N° d'ordre: 9691

UNIVERSITÉ PARIS-SUD
FACULTÉ DES SCIENCES D'ORSAY

THÈSE

Présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS XI

Spécialité : Mathématiques

par

Amandine VÉBER

THÉORÈMES LIMITES POUR DES PROCESSUS DE BRANCHEMENT ET DE COALESCENCE SPATIAUX

Rapporteurs :
Prof. Jean BERTOIN
Prof. Anton WAKOLBINGER

Soutenue le 08 décembre 2009 devant la commission d'examen :

Prof. Alison ETHERIDGE	(Directrice de thèse)
Prof. Jean-François LE GALL	(Directeur de thèse)
Prof. Sylvie MÉLÉARD	(Présidente du jury)
Prof. Alain ROUAULT	(Examinateur)
Prof. Anton WAKOLBINGER	(Rapporteur)

Remerciements

Après trois ans passés sur le chemin des arts mathématiques, il est temps de se retourner un moment pour saluer ses maîtres et compagnons de route...

Mes premiers remerciements vont à mes directeurs, Alison Etheridge et Jean-François Le Gall. Grâce à eux j'ai appris la rigueur et l'intuition, profité d'une certaine autonomie tout comme de leur disponibilité et découvert de nouveaux horizons tant mathématiques que géographiques.

Merci à Jean Bertoin et Anton Wakolbinger de m'avoir fait l'honneur de rapporter ma thèse et à Sylvie Méléard et Alain Rouault celui d'avoir accepté de faire partie de mon jury.

J'ai eu le privilège de croiser la route de Julien et Nathanaël Berestycki, Louigi Addario-Berry, Christina Goldschmidt, James Martin, Raazesh Sainudiin, Jay Taylor et Feng Yu qui, pour une raison que cette thèse n'éclaircira certainement pas, m'ont offert de travailler avec eux. Je les remercie aussi sincèrement pour leurs conseils et leur amitié inestimables (ainsi que pour un splendide séjour dans une île ensoleillée dont je n'ai plus le droit de prononcer le nom...).

Des amis, le département de mathématiques de l'E.n.s. et celui de statistiques d'Oxford n'en manquent pas. Puisque je ne peux pas citer tout le monde, qu'il soit quand même rendu justice à Vincent Rivoirard et Marie Théret pour m'avoir chaleureusement accueillie au V6, supportée depuis, et avoir fait de ce bureau un réservoir d'idées et une palette de souvenirs mémorables. Les autres se reconnaîtront entre les lignes, qui sont assez espacées pour contenir également les anciens occupants du passage vert Jérémie Bettinelli, Laurent Ménard, Grégory Miermont, Patricia Reynaud-Bouret et Mathilde Weill et ceux qui y passent encore de temps en temps. Même si j'ai moins profité de l'ambiance orsayenne, je n'en oublie pas toutes les personnes que j'y ai croisées et les remercie pour l'aide souvent discrète mais indéniable que ces contacts m'ont apportée.

Pour finir, mention spéciale à ma famille et à tous mes amis, que j'aime assez pour qu'ils se reconnaissent simplement aux sourires que leur présence fait naître sur mon visage. Robert et Sophie, merci du plus profond de mon coeur pour tout ce que vous m'offrez, parfois sans même vous en rendre compte. Quant à mes chers parents et à Angéline et Florent, ce travail vous est entièrement dédié.

Théorèmes limites pour des processus de branchement et de coalescence spatiaux

Résumé : Dans une première partie, on considère une famille indexée par $\varepsilon \in (0, 1/2)$ de super-mouvements browniens (resp., de mouvements browniens branchants) critiques dont les particules sont tuées instantanément (resp., à taux ε) dans une collection Γ_ε d'obstacles distribués dans $\mathbb{R}^d$ suivant un processus ponctuel de Poisson. Dans le cas des super-mouvements browniens, Γ_ε est constitué de boules de rayon ε dont l'intensité est proportionnelle à $\log(\varepsilon^{-1})$ si $d = 2$ et ε^{2-d} si $d \geq 3$, tandis que la configuration d'obstacles ne dépend pas du paramètre ε pour les diffusions branchantes. On montre alors que, presque sûrement pour la loi des obstacles, la loi des processus de branchement tués dans Γ_ε (éventuellement renormalisés) converge lorsque ε tend vers 0 vers celle d'un superprocessus dont les particules sont tuées de manière homogène dans $\mathbb{R}^d$. On en déduit des propriétés d'atteinte du complémentaire d'un très grand domaine par des mouvements browniens branchants faiblement tués. Dans une deuxième partie, on s'intéresse aux généalogies d'individus correspondant à des modèles particuliers de populations spatiales. Dans le cas du modèle d'îles de Wright avec D communautés, on généralise des résultats de Wakeley en donnant des conditions nécessaires et suffisantes pour que le processus généalogique d'un échantillon fini d'individus converge en loi vers un coalescent à collisions multiples et simultanées non-spatial lorsque D tend vers l'infini. On définit ensuite le processus Λ-Fleming-Viot spatial et on en étudie les généalogies lorsque la population ainsi modélisée est répartie sur un tore (en dimension 2) de côté L et est soumise à des événements de reproduction d'échelles spatiales et temporelles très différentes. On établit alors les différentes limites possibles des processus généalogiques associés lorsque L tend vers l'infini, en fonction de la rareté et de la portée des *grands* événements.

Mots-clés : processus de branchement spatiaux, coalescence, homogénéisation, convergence.

Limit theorems for spatial branching and coalescent processes

Abstract: In a first part, we consider a family of critical super-Brownian motions (resp., branching Brownian motions), indexed by $\varepsilon \in (0, 1/2)$, whose particles are killed instantaneously (resp., at rate ε) inside a collection Γ_ε of obstacles distributed according to a Poisson point process on $\mathbb{R}^d$. In the case of super-Brownian motion, Γ_ε is made of balls of radius ε and the intensity of the point process is proportional to $\log(\varepsilon^{-1})$ if $d = 2$ and ε^{2-d} if $d \geq 3$, whereas the obstacle configuration does not depend on ε for the branching diffusions. We show that, almost surely for the law of the obstacles, the laws of the branching processes killed within Γ_ε (and possibly rescaled) converge as ε tends to zero to that of a superprocess whose particles are killed homogeneously in $\mathbb{R}^d$. Elaborating on this result, we establish hitting probabilities of the complement of a very large domain by weakly killed branching Brownian motion. In a second part, we are interested in genealogies of individuals corresponding to particular spatial population models. In the case of Wright's island model with D demes, we generalize results of Wakeley by giving necessary and sufficient conditions for the genealogical process of a finite sample of individuals to converge in distribution to a non-spatial coalescent with simultaneous multiple mergers, as D tends to infinity. We next define the spatial Λ-Fleming-Viot process and study its genealogies when the population is distributed on a two-dimensional torus of side L and is subject to reproduction events of different time- and space-scales. We then establish the different possible limits for the corresponding genealogical processes as L tends to infinity, according to the frequence and the extent of the *large* events.

Keywords: spatial branching processes, coalescence, homogenization, convergence.
AMS Classification: 60J80, 60K37, 60B10, 60J25, 92D10.

Erratum

In Chapter 5, the results of Theorem 5.15 are not entirely accurate. As explained below, when $\alpha = 1$ *but* $\psi_L \ll L^2$*, the limiting behaviour of the genealogical process* $\mathcal{A}^L$ *depends on the precise relations between* ρ_L *and* ψ_L^2*, and between* ρ_L *and* L^2*. When* $\psi_L = cL$ *for a constant* $c > 0$*, the limits given in Theorem 5.15 have essentially the right form, but depend also on c. Hence, the end of Section 3 of Chapter 5 should be replaced from the paragraph after Theorem 5.11 on by the following text, and the required (minor) modifications of the proofs can be found in* [**BEV09**].

For $\alpha = 1$, things are more complicated. When ψ_L is commensurate with L, large scale events cover a non-negligible fraction of the torus. If they happen too quickly, then they will be able to capture multiple lineages while the locations of those lineages are still correlated with their starting points. For intermediate ranges of ρ_L, lineages will have homogenised their positions on $\mathbb{T}(L)$ through small events, but not coalesced, before the first large event occurs and we can expect a Λ-coalescent limit. If the large events are too rare, then coalescence will be through small events and we shall recover the Kingman coalescent again.

To give a precise result we need to define the limiting objects that arise. In the case $\alpha = 1$, for each $L \in \mathbb{N}$, we set

$$\varpi_L = \begin{cases} \rho_L & \text{if } \rho_L/(L^2 \log L) \text{ has a finite limit,} \\ \frac{L^2 \log L}{2\pi\sigma_s^2} & \text{if } \rho_L/(L^2\log L) \to +\infty, \end{cases}$$

and define $\mathcal{A}^{L,u}$ as before. Since we shall need to keep track of the labels (spatial positions) of the ancestral lineages in some cases, it will also be convenient to introduce the following rescaling of $\mathcal{A}^L$, evolving on $\mathbb{T}(1)$ for all $L \in \mathbb{N}$:

$$\bar{\mathcal{A}}^L(t) = \frac{1}{L}\,\mathcal{A}^L(\varpi_L t), \qquad t \geq 0,$$

where by this notation we mean that the labels are rescaled by a factor L^{-1}. Similarly, for $\mathbf{x} \in \mathbb{T}(1)^n$ we write $L\mathbf{x}$ for $(Lx_1, \dots, Lx_n) \in \mathbb{T}(L)^n$. Finally, let us introduce the processes which will appear as the limits of our rescaled genealogical processes.

DEFINITION 5.12. *Let* $b \in [0,\infty)$ *and* $c > 0$*. We call* $\bar{\mathcal{A}}^{\infty,b,c}$ *the Markov process with values in* $\bigcup_{n\in\mathbb{N}} \mathcal{P}_n^\ell$ *(with labels in* $\mathbb{T}(1)$*) such that*

1. *The labels of the lineages perform independent Brownian motions on* $\mathbb{T}(1)$ *at speed* $b\sigma_s^2$ *(if* $b = 0$*, the labels are constant), until the first large event occurs.*
2. *Large events are generated by a Poisson point process* $\overline{\Pi}^B$ *on* $\mathbb{R} \times \mathbb{T}(1/c) \times (0, 1/\sqrt{2}]$ *with intensity measure* $dt \otimes dx \otimes \mu^B(dr)$*. At a point* (t,x,r) *of* $\overline{\Pi}^B$*, a number* $u \in [0,1]$ *is sampled from the probability measure* ν_r^B*, and each lineage whose label belongs to* $B_{\mathbb{T}(1)}(cx, cr)$ *is affected (resp., is not affected) by the event with probability* u *(resp.,* $1-u$*), independently of each other. A label* z *is chosen uniformly at random in* $B_{\mathbb{T}(1)}(cx,cr)$*, and all the lineages affected merge into one block which adopts the label* z*. The other lineages (blocks and labels) remain unchanged.*
3. *The evolution of the labels starts again in the same manner.*

REMARK 5.13. *Notice that this process looks like another spatial* Λ*-coalescent, except that now ancestral lineages perform independent spatial motions in between coalescence events. This process is dual (in the obvious way) to a spatial* Λ*-Fleming-Viot process in which, during their lifetimes, individuals move around in space according to independent Brownian motions.*

For each $r \in [0, 1/\sqrt{2}]$, let V_r denote the volume of the ball $B_{\mathbb{T}(1)}(0,r)$.

DEFINITION 5.14. *Let* $\beta \in [0,\infty)$ *and* $c > 0$. *We use* $\Lambda^{(\beta,c)}$ *to denote the* Λ*-coalescent, defined on* $\bigcup_{n\in\mathbb{N}} \mathcal{P}_n$, *for which if there are currently* m *ancestral blocks, then each transition involving* k *of them merging into one happens at rate*

$$\lambda^{(\beta,c)}_{m,k} = c^{-2}\int_0^{(\sqrt{2})^{-1}}\int_0^1 (V_{cr}u)^k(1-V_{cr}u)^{m-k}\nu_r^B(du)\mu^B(dr) + \beta\,\delta_{\{k=2\}}.$$

Recall the notation $\wp_n$ and $\wp_n(\mathbf{x})$ introduced in Notation 5.6, and $\mathcal{L}_{\mathrm{P}}(X)$ and $\Rightarrow$ introduced in the statement of Theorem 5.11. We can now state the result for $\alpha = 1$.

THEOREM 5.15. *Suppose there exists* $c > 0$ *such that for every* $L \in \mathbb{N}$, $\psi_L = cL$. *Let* $n \in \mathbb{N}$, $\mathbf{x} \in \mathbb{T}(1)^n$ *such that* $x_i \neq x_j$ *whenever* $i \neq j$, *and let* $(A_L)_{L\in\mathbb{N}}$ *be such that for every* L, $A_L \in \Gamma_{\mathcal{A}}(L,n)^*$. *Then, as* $L \to \infty$,
(a) *If* $\rho_L L^{-2} \to b \in [0,\infty)$,

$$\mathcal{L}_{\mathbb{P}_{\wp_n(L\mathbf{x})}}\big(\bar{\mathcal{A}}^L\big) \Rightarrow \mathcal{L}_{\mathbb{P}_{\wp_n(\mathbf{x})}}\big(\bar{\mathcal{A}}^{\infty,b,c}\big),$$

(b) *If* $\rho_L L^{-2} \to \infty$, $\frac{2\pi\sigma_s^2\rho_L}{L^2\log L} \to \beta \in [0,\infty)$ *and if the total rate of occurrence of large events is finite (i.e.,* μ^B *has finite total mass),*

$$\mathcal{L}_{\mathbb{P}_{A_L}}\big(\mathcal{A}^{L,u}\big) \Rightarrow \mathcal{L}_{\mathbb{P}_{\wp_n}}\big(\Lambda^{(\beta,c)}\big).$$

(c) *If* $\frac{\rho_L}{L^2\log L} \to \infty$,

$$\mathcal{L}_{\mathbb{P}_{A_L}}\big(\mathcal{A}^{L,u}\big) \Rightarrow \mathcal{L}_{\mathbb{P}_{\wp_n}}\big(\mathcal{K}\big).$$

Notice that the case (a) differs from all other cases in that the influence of space does not disappear as $L \to \infty$ and the evolution of the limiting genealogy still depends on the precise locations of the lineages.

The intuition behind Theorem 5.15 is as follows. If $\psi_L \propto L$ large events cover a non-negligible fraction of the torus, and so only a few large events are sufficient to gather two lineages at a distance at which they can coalesce. However, a local central limit theorem will give us that on a timescale of order at most $\mathcal{O}(L^2)$, a lineage subject to only small events behaves approximately like Brownian motion, whereas after a time $t_L \gg L^2$, its distribution is nearly uniform on $\mathbb{T}(L)$ (for L large enough, see Lemma 5.23). Since the mean time before a large event affects a lineage is of order $\mathcal{O}(\rho_L)$, the limiting genealogical process (when we include both large and small reproduction events) will depend on how ρ_L scales with L^2. If ρ_L is of order at most $\mathcal{O}(L^2)$, then space matters and the process $\mathcal{A}^L$ rescaled to evolve on $\mathbb{T}(1)$ on the timescale ρ_L converges to a system of coalescing Brownian motions, whereas if $\rho_L \gg L^2$, the homogenisation of the labels/locations of the lineages before the occurrence of the first large event which affects them leads to a limiting unlabelled genealogical process given by an exchangeable coalescent with multiple mergers.

REMARK 5.16. *It is somehow disappointing that we must impose a finite rate of large events to obtain the convergence of Theorem 5.15(b). Indeed, it seems that case (a) should give us the right picture: in the limit, in between large events lineages perform Brownian motions on the torus of sidelength 1 due to small events, except that now the time required for at least one lineage to be affected by a large event is so long that lineages exhaust space and their locations become uniformly distributed over the torus before they are taken by a coalescence event. However, when* μ^B *has infinite mass, lineages are infinitely often in the (geographical) range of a large reproduction event over any interval of time, and we need good control of their complete paths to actually be able to say something about the epoch and outcome of the first potential coalescence event. Now, observe that Equation (5.65) can only be generalized to the finite-dimensional distributions* *of these paths, and does not guarantee that a large event cannot capture some of the lineages at a time when they are not uniformly distributed over* $\mathbb{T}(1)$.

Theorem 5.15 deals with the case where ψ_L is proportional to L. Let us now comment on the remaining cases, in which $\alpha = 1$ but $\psi_L \ll L$. First, it is easy to see that the convergence in (c) still holds, since it is based on the fact that large events are so rare that none of them occurs before small events reduce the genealogical process to a single lineage.

Second, since the total rate of large events on the timescale ρ_L is $\mu_B(\mathbb{R}_+)L^2/\psi_L^2$, it cannot be bounded unless $\mu^B = 0$ (a situation we excluded). On the other hand, for the reason expounded in Remark 5.16 we are unable to derive a limiting behaviour for the genealogy when large events can accumulate, and so the result of Theorem 5.15(b) has no counterpart when $\psi_L \ll L$.

Third, as explained above, when $\rho_L \leq bL^2$ any limiting process will necessarily have a spatial component. Now, because we start with lineages at distance $\mathcal{O}(L)$ of each other, we need to rescale space by L in order to obtain a non trivial initial condition. The last parameter we need is the timescale ϖ_L on which to consider the genealogical process. But a separation of timescales will not occur here, and so the computations done in Section 5 will show that the suitable choice of ϖ_L depends on the precise behaviour of ρ_L/L^2 and ρ_L/ψ_L^2. Several limiting processes are thus possible, and since all the arguments needed to derive these limits are scattered in Sections 5 and 7, we chose not to detail them here.

Contents

CHAPITRE 1

Introduction

Dans cette thèse, on s'intéresse à deux types de questions : le comportement d'un processus de branchement spatial évoluant parmi des obstacles aléatoires et la description de la généalogie d'un échantillon d'individus correspondant à des modèles particuliers de populations réparties dans un espace.

Contents

1. Processus de branchement spatiaux en milieu aléatoire

Les processus de branchement sont des modèles de population (au sens large) à la fois assez simples pour que l'on puisse en étudier le comportement au cours du temps avec une chance raisonnable de succès (on verra par exemple que la *propriété de branchement* qui leur est commune produit une structure agréable) et assez riches pour en tirer des évolutions très variées. Les plus connus sont les processus de Galton-Watson à temps discret, qui décrivent la taille d'une population dans laquelle chaque individu de la génération présente produit indépendamment des autres un nombre aléatoire de descendants. La génération suivante est alors constituée de l'ensemble de ces descendants, dont le nombre total devient la nouvelle valeur du processus de Galton-Watson associé.

Une fois l'évolution de ces modèles bien comprise, une généralisation naturelle est de supposer les individus sous-jacents répartis dans un espace, dans lequel ils continuent à se reproduire indépendamment les uns des autres mais peuvent aussi se déplacer. Nous nous intéresserons à deux types particuliers de ces modèles spatiaux : le *processus de branchement brownien* et le *super-mouvement brownien* que nous introduirons dans le paragraphe 1.2.1. Outre son interprétation en termes de collection de particules évoluant et se reproduisant dans un espace, d'autres intérêts du super-mouvement brownien résident dans ses nombreuses connexions avec certaines équations aux dérivées partielles et avec des modèles de particules en interaction (modèle du votant, processus de contact) ou de mécanique statistique. On renvoie à l'introduction de [**LG99**] pour un aperçu de ces liens.

Commençons donc par rappeler quelques résultats sur les processus de branchements non-spatiaux, à espace d'états discret ou continu. Ces propriétés constituent d'une part une bonne introduction aux différentes caractérisations des processus de branchement spatiaux et aux liens entre ces derniers ; d'autre part, elles nous seront utiles lorsque nous aurons introduit une composante spatiale aux populations que l'on veut étudier.

1.1. Processus de branchement non-spatiaux. On s'intéressera exclusivement aux processus de branchement à temps continu (dans lesquels une seule particule se reproduit à la fois), bien que la grande majorité des résultats cités pour les processus de Galton-Watson soient également valables pour leurs équivalents à temps discret. Informellement, un *processus de Galton-Watson à temps continu* représente la taille d'une population dans laquelle chaque particule indépendamment des autres vit pendant un temps exponentiellement distribué de paramètre $\rho > 0$, fixé et commun à toutes les particules, puis « branche », i.e., meurt et est remplacée par un nombre aléatoire de descendants qui suivent ensuite la même évolution. Ce nombre, qui peut être égal à 0 (auquel cas la particule disparue n'est simplement pas remplacée), est donné par une réalisation indépendante d'une variable aléatoire K à valeurs dans $\mathbb{N} \cup \{0\}$ (par convention, on note $\mathbb{N}$ l'ensemble des entiers strictement positifs). Notons Φ la fonction génératrice de K.

Grâce à la propriété d'oubli des variables exponentielles, on obtient alors un processus markovien. L'indépendance d'évolution entre les particules garantit quant à elle que ce processus possède la *propriété d'additivité* (ou *de branchement*) suivante :

PROPOSITION 1.1. (Propriété d'additivité) *Soient* $N^{(n_1)}$ *et* $N^{(n_2)}$ *deux processus de Galton-Watson indépendants, de paramètres* ρ *et* Φ *identiques et de valeurs initiales respectives* $n_1, n_2 \in \mathbb{N}$. *Alors,*

$$N^{(n_1)} + N^{(n_2)} \overset{(d)}{=} N^{(n_1+n_2)},$$

où $N^{(n_1+n_2)}$ *est un processus de Galton-Watson de paramètres* ρ *et* Φ, *et de valeur initiale* n_1+n_2.

Ainsi, une somme finie de Galton-Watson indépendants et de mêmes caractéristiques est également un Galton-Watson de valeur initiale la somme des toutes les valeurs initiales des processus ajoutés.

Une première question naturelle est de savoir si un tel processus est bien défini à tout temps, i.e., si la taille de la population n'explose pas en temps fini. Le Théorème III.2.1 de [**AN72**] assure que N_t est fini pour tout $t \geq 0$ p.s. si et seulement si pour tout $\varepsilon > 0$,

$$\int_{1-\varepsilon}^{1} \frac{ds}{\Phi(s) - s} = \infty. \tag{1.1}$$

Dans le cas où $\mathbb{E}[K] = \Phi'(1) < \infty$, la condition (1.1) est satisfaite et N prend donc ses valeurs dans $\mathbb{N} \cup \{0\}$ à tout temps.

Comme on peut le constater, la fonction génératrice Φ joue un rôle important dans l'analyse du comportement de N. On renvoie au chapitre III de [**AN72**] pour une preuve des résultats cités ci-dessous et de nombreux compléments. Rappelons qu'un processus de Galton-Watson est appelé *critique* (resp., *sous-critique*, *surcritique*) si $\Phi'(1) = 1$ (resp., $\Phi'(1) < 1$, $\Phi'(1) \in (1, \infty]$). Le nombre moyen de descendants d'une particule dicte le comportement asymptotique du processus : celui-ci atteint la valeur absorbante 0 (il « s'éteint ») en temps fini p.s. s'il est critique ou sous-critique (on exclut bien sûr le cas trivial où $K = 1$ p.s.) ; dans le cas surcritique, il croît exponentiellement avec une probabilité strictement positive et s'éteint sinon. Par ailleurs, si on fixe $\lambda \in [0,1]$ et que l'on définit pour tout $t \geq 0$

$$v(t,\lambda) := \mathbb{E}\big[\lambda^{N_t} \big| N_0 = 1\big],$$

la fonction v satisfait l'équation intégrale

$$v(t,\lambda) - \rho \int_0^t \big(\Phi(v(s,\lambda)) - v(s,\lambda)\big) ds = \lambda. \tag{1.2}$$

Ce type de caractérisation apparaîtra pour chacun des processus que l'on verra dans la suite.

Passons à un autre type de processus de branchement, dont l'espace d'états est cette fois $\mathbb{R}_+$ tout entier. Après avoir donné la définition des *processus de branchement à espace d'états continu* et quelques propriétés, nous reviendrons sur les relations entre ces derniers et les processus de Galton-Watson.

DÉFINITION 1.2. *Un processus markovien $\{Y_t,\ t \geq 0\}$ à valeurs dans $\mathbb{R}_+$ est un processus de branchement à espace d'états continu (ou CSBP d'après la terminologie anglaise) si ses noyaux de transition $P_t(x,dy)$ vérifient la* propriété d'additivité *suivante : pour tout $t \geq 0$ et $x, x' \in \mathbb{R}_+$,*

$$P_t(x,\cdot) * P_t(x',\cdot) = P_t(x+x',\cdot),$$

où $*$ représente le produit de convolution de deux mesures.

On aura bien sûr reconnu dans cette définition l'équivalent de la propriété d'additivité des processus de Galton-Watson, qui est le dénominateur commun de tous les processus de branchement. A nouveau, on appellera un CSBP Y *critique* (resp., *sous-critique*, *surcritique*) si pour tout $x \in \mathbb{R}_+$ et $t \geq 0$, $\mathbb{E}[Y_t | Y_0 = x] = x$ (resp., $< x$, $> x$). Les CSBP critiques et sous-critiques ont une caractérisation très utile, qui découle en particulier de la propriété d'additivité et du fait qu'ils n'explosent pas en temps fini.

THÉORÈME 1.3. *Soit Y un CSBP critique ou sous-critique et non nul. Alors sa transformée de Laplace doit être de la forme*

$$\int P_t(x,dy)\, e^{-\lambda y} = e^{-x u(t,\lambda)}, \qquad \lambda \geq 0, \tag{1.3}$$

où la fonction $u(t,\lambda)$ est l'unique solution positive de l'équation intégrale

$$u(t,\lambda) + \int_0^t \psi\big(u(s,\lambda)\big) ds = \lambda, \tag{1.4}$$

pour une fonction ψ de la forme

$$\psi(u) = \alpha u + \beta u^2 + \int_0^\infty \big(e^{-ru} - 1 + ru\big) \pi(dr), \tag{1.5}$$

où $\alpha \geq 0$, $\beta \geq 0$ et π est une mesure σ-finie sur $(0,\infty)$ telle que $\int_0^\infty (r \wedge r^2) \pi(dr) < \infty$.

Réciproquement, si ψ est une fonction de ce type, alors il existe un unique (en loi) CSBP associé à ψ. La fonction ψ est appelée mécanisme de branchement *de Y.*

On peut trouver ce résultat et sa preuve complète dans [**LG99**] (Théorème II.1). Dans la définition de ψ, le paramètre α correspond à un taux de meurtre : si $\beta = 0$ et π est la mesure nulle, on peut résoudre explicitement l'équation intégrale correpondante et obtenir ainsi $Y_t = Y_0 e^{-\alpha t}$ pour tout $t \geq 0$. Le paramètre β donne la vitesse de la composante diffusive de Y : si $\alpha = 0$ et

$\pi = 0$, on obtient cette fois

$$u(t,\lambda) = \frac{\lambda}{1+\beta t\lambda}, \qquad t,\lambda \geq 0,$$

qui correspond à une diffusion que l'on appelle *diffusion de Feller* de paramètre β. Pour relier β à la vitesse de la diffusion, observons que cette dernière peut également être construite comme la solution de l'équation différentielle stochastique

$$dY_t = \sqrt{2\beta Y_t}\, dB_t,$$

où B est un mouvement brownien standard unidimensionnel. Enfin, la mesure π décrit les sauts (toujours positifs) de Y.

Explorons un moment les propriétés d'un tel objet. La caractérisation sous forme de transformée de Laplace permet de voir que le semi-groupe de transition associé à un CSBP possède la propriété de Feller (voir par exemple [**EK86**], p.166, pour une définition), et par conséquent d'après le Théorème 4.2.7 de [**EK86**] un CSBP possède une modification p.s. càdlàg et satisfait la propriété de Markov forte. Observons d'ailleurs que, puisqu'il est à valeurs positives, par définition un CSBP critique (resp., sous-critique) est une martingale (resp., une surmartingale). Un problème d'intégrabilité peut se poser dans le cas surcritique. Par analogie avec les processus de Galton-Watson, on peut poser la question de l'extinction d'un CSBP. Cette fois, c'est un critère intégral impliquant ψ qui en donne la réponse. Grey a montré dans [**Gr74**] qu'un CSBP s'éteint en temps fini presque-sûrement si et seulement si son mécanisme de branchement ψ satisfait

$$\int^{\infty} \frac{du}{\psi(u)} < \infty.$$

Si cette condition n'est pas remplie, le temps d'extinction est alors infini p.s. On peut montrer par des arguments de martingales qu'un CSBP (sous-)critique converge vers 0 p.s., mais en reprenant le cas $\beta = 0, \pi = 0$ on voit qu'il ne s'éteint pas nécessairement en temps fini. On pourra interpréter ce fait plus tard comme étant la marque du caractère infini de la population sous-jacente.

Pour clore ce paragraphe, intéressons-nous brièvement aux relations entre Galton-Watson et CSBP. S'il est vrai que tout CSBP peut être obtenu comme limite faible d'une suite de processus de Galton-Watson bien choisis (cf. [**Lam67**]), nous allons concentrer notre attention sur l'une des manières d'obtenir une diffusion de Feller, car nous reverrons les renormalisations utilisées ici dès le paragraphe suivant. Considérons un processus de Galton-Watson critique N, de taux de reproduction $\rho > 0$ et de fonction génératrice du nombre de descendants Φ. On a donc $\Phi'(1) = 1$ et supposons en outre que $\Phi''(1) < \infty$ (notons que $\Phi''(1)$ est aussi la variance du nombre de descendants d'une particule, que l'on note de manière plus standard σ^2). Notre objectif est de faire tendre la taille de la population vers l'infini, tout en renormalisant N pour faire en sorte que le processus limite reste à valeurs finies. C'est une convergence en loi que nous cherchons à obtenir ; donnons-nous donc une suite $(N^{[n]})_{n\in\mathbb{N}}$ de Galton-Watson de caractéristiques ρ et Φ et supposons que la suite de variables réelles $(N^{[n]}(0)/n)_{n\in\mathbb{N}}$ converge en loi vers une variable $Y_{(0)}$. Mesurons alors la taille de la population par unités de n particules et le temps également par périodes de n unités initiales de temps, i.e., pour tout $t \geq 0$ posons

$$N_t^n := \frac{1}{n} N_{nt}^{[n]}.$$

Feller a montré dans [**Fel51**] que lorsque n tend vers $+\infty$, la suite de processus $(N^n)_{n\in\mathbb{N}}$ converge au sens de la convergence en loi des marginales de dimension finie vers une diffusion de Feller de paramètre $\rho\sigma^2/2$ et dont la valeur initiale est distribuée selon la loi de $Y_{(0)}$. Pour rappel, la convergence au sens de celle des marginales de dimension finie d'une suite de processus aléatoires

$(X^n)_{n\in\mathbb{N}}$ vers un processus X^∞ signifie que pour tout $k \in \mathbb{N}$ et $0 \leq t_1 < \cdots < t_k < \infty$, les vecteurs $(X^n_{t_1}, \ldots, X^n_{t_k})$ convergent en loi vers le vecteur $(X^\infty_{t_1}, \ldots, X^\infty_{t_k})$. Donnons quelques arguments heuristiques pour se convaincre qu'un tel résultat est raisonnable. Tout d'abord, si la taille de la population initiale à la n-ième étape est de l'ordre de n, multiplier $N^{[n]}$ par $1/n$ fait en sorte que celle-ci ne devienne pas arbitrairement grande (au moins pour les temps proches de 0 a priori). Par ailleurs, rendre les variations infinitésimales permet d'obtenir à la limite un processus à trajectoires continues à valeurs dans $\mathbb{R}_+$. Pour comprendre le changement d'échelle temporelle, considérons d'abord une marche aléatoire sur $\mathbb{Z}$ dont les pas ont une espérance nulle et une variance finie et renormalisons-la aussi par $1/n$. Au premier ordre, les pas vers la gauche et vers la droite se compensent et par conséquent il faut un temps de l'ordre de n^2 pour que la variance des sauts se fasse sentir et que le processus renormalisé atteigne une distance non-triviale. On obtient ainsi la renormalisation conduisant au mouvement brownien. Dans le cas du processus de Galton-Watson critique, puisque $\Phi'(1) = 1$, la disparition de certaines particules non-remplacées compense en moyenne la production de nouvelles particules et donc d'après l'analogie avec la marche aléatoire, on s'attend à avoir besoin d'un temps d'ordre n^2 pour observer une variation significative de la taille de la population. Cependant, il faut se rappeler que celle-ci est constituée de $\mathcal{O}(n)$ particules, chacune essayant de brancher indépendamment à taux ρ : les événements de branchement arrivent donc déjà à taux $\mathcal{O}(n)$ avant de renormaliser le temps, et c'est une période de l'ordre de n unités de temps qu'il faut attendre pour que $N^{[n]}/n$ varie de façon macroscopique. On renvoie aux paragraphes correspondants sur les convergences de processus spatiaux pour des arguments plus rigoureux.

Un mot de philosophie pour finir. L'interprétation en termes de taille de population semble plus hasardeuse pour un CSBP du fait qu'il prend des valeurs non entières. Néanmoins, sa construction comme limite de processus de Galton-Watson nous incite à le considérer comme la « masse » totale d'une population infinie, dans laquelle chaque particule a une masse infinitésimale. On pourra garder cette image lorsqu'on rajoutera une composante spatiale à nos populations : chaque « particule » d'un super-mouvement brownien donnera un poids infinitésimal au point de l'espace où elle se trouve.

1.2. Mouvement brownien branchant et super-mouvement brownien. Introduisons à présent les différents processus de branchement spatiaux que nous étudierons par la suite. Le premier décrit une collection de particules se déplaçant dans $\mathbb{R}^d$ et se reproduisant à taux fini au cours du temps, le tout indépendamment les unes des autres. Le second peut être vu comme une limite du premier, dans lequel on fait tendre le nombre de particules et leur taux de reproduction vers l'infini de façon appropriée. Mis à part cette description informelle de leurs mécanismes d'évolution, ces deux types de processus ont en commun d'être à valeurs dans l'espace $\mathcal{M}_f(\mathbb{R}^d)$ des mesures finies sur $\mathbb{R}^d$: si $\mu \in \mathcal{M}_f(\mathbb{R}^d)$ est l'état présent de notre population, alors $\mu(A)$ donne la masse de particules présentes dans la région $A \subset \mathbb{R}^d$. On utilisera les notations $\langle \mu, f\rangle$ pour l'intégrale de la fonction f contre la mesure μ et $\mathcal{B}_b^+(\mathbb{R}^d)$ (resp., $C_b^+(\mathbb{R}^d)$) pour l'ensemble des fonctions mesurables (resp., continues), positives et bornées sur $\mathbb{R}^d$.

1.2.1. *Définitions et caractérisation via des équations d'évolution.* On choisit ici comme déplacement des particules un mouvement brownien dans $\mathbb{R}^d$ ($d \geq 1$), mais les objets suivants peuvent être définis de manière plus générale en utilisant un processus markovien sur un espace polonais.

Fixons $n \in \mathbb{N}$, $\rho > 0$ et écrivons $(\xi_1(0), \ldots, \xi_n(0))$ les positions au temps 0 des n particules qui composent la population initiale. Au cours du temps, ces particules évoluent de la manière suivante : chacune, indépendamment des autres,

- se déplace dans $\mathbb{R}^d$ selon la loi d'un mouvement brownien,
- branche à taux ρ et laisse à l'endroit où elle se trouvait un nombre aléatoire de descendants. Ce nombre est donné par une réalisation indépendante d'une variable aléatoire K à valeurs dans $\mathbb{N} \cup \{0\}$, dont on notera Φ la fonction génératrice.

Chaque particule nouvellement créée est indexée de manière cohérente (nous n'insistons pas sur ce point, qui n'a pas d'importance pour la suite) et évolue ensuite de la même manière, indépendamment de ses « soeurs » et du reste de la population. On notera $I(t)$ l'ensemble des indices des particules en vie à l'instant $t \geq 0$. Ceci définit un système de particules, que l'on représente par sa mesure empirique : pour tout $t \geq 0$, on pose

$$Z_t := \sum_{i \in I(t)} \delta_{\xi_i(t)}, \tag{1.6}$$

où δ_x est la masse de Dirac en $x \in \mathbb{R}^d$ et $\xi_i(t) \in \mathbb{R}^d$ est la position de la particule i à l'instant t. Le processus aléatoire $\{Z_t,\ t \geq 0\}$ que l'on obtient ainsi est appelé *processus de branchement brownien* ou *mouvement brownien branchant*. Comme annoncé, il est bien à valeurs dans les mesures positives (ponctuelles, en fait) sur $\mathbb{R}^d$ et son caractère markovien découle de la propriété d'oubli des variables exponentielles dictant le temps de vie de chaque particule.

Rien ne dit cependant que Z reste à valeurs dans les mesures finies à tout instant. Pour le voir, il suffit d'observer que la masse totale de la population, $\langle Z, 1\rangle$, suit la loi d'un processus de Galton-Watson en temps continu de paramètres ρ et Φ. Le résultat dépend donc du critère énoncé dans (1.1). Par ailleurs, ce lien avec les processus de Galton-Watson conduit naturellement à la terminologie suivante : on appellera un mouvement brownien branchant *critique* (resp., *surcritique*, *sous-critique*) si $\mathbb{E}[K] = 1$ (resp., $\mathbb{E}[K] > 1$, $\mathbb{E}[K] < 1$). Dans ce qui suit, on ne considèrera que des processus critiques ou sous-critiques. Par conséquent, d'une part (1.1) est satisfait, et d'autre part Z s'éteint en temps fini p.s. dès lors que sa valeur initiale a une masse totale finie.

Toujours par analogie avec le processus de Galton-Watson sous-jacent, il est facile de voir qu'un processus de branchement brownien possède *la propriété de branchement.*

PROPOSITION 1.4. (Propriété de branchement) *Soient $Z^{(\mu_1)}$ et $Z^{(\mu_2)}$ deux mouvements browniens branchants indépendants de même mécanisme de branchement et de valeurs initiales respectives $\mu_1, \mu_2 \in \mathcal{M}_p(\mathbb{R}^d)$ (l'ensemble des mesures ponctuelles sur $\mathbb{R}^d$). Alors*

$$Z^{(\mu_1)} + Z^{(\mu_2)} \overset{(d)}{=} Z^{(\mu_1+\mu_2)}, \tag{1.7}$$

où $Z^{(\mu_1+\mu_2)}$ est un mouvement brownien branchant de même mécanisme de branchement et de valeur initiale $\mu_1 + \mu_2$.

Lue de gauche à droite, l'égalité en loi (1.7) nous dit que si l'on superpose des mouvements browniens branchants indépendants et de mêmes caractéristiques, on obtient à nouveau un processus de branchement brownien. Lue de droite à gauche, elle montre que l'on peut décomposer Z en une somme de sous-familles (par exemple celles produites par chacune des particules initiales) qui évoluent indépendamment les unes des autres.

Comme pour les processus de branchement non-spatiaux, cette propriété importante est à l'origine de la caractérisation de Z par des *équations d'évolution*. Parce que nous aurons besoin d'une forme assez générale dans les chapitres suivants, oublions pour un temps que le mouvement des particules est donné par un mouvement brownien et notons ξ le processus markovien (à valeurs dans un espace F, qui sera simplement $\mathbb{R}^d$ pour nos applications) décrivant le déplacement de l'une d'entre elles. Notons aussi

- P_x la mesure de probabilité sur l'espace de définition de ξ sous laquelle $\xi_0 = x \in F$ p.s. et E_x l'espérance associée ;
- $\mathbb{P}_\mu$ la mesure de probabilité sur l'espace de définition de Z sous laquelle $Z_0 = \mu \in \mathcal{M}_p(F)$ p.s. et $\mathbb{E}_\mu$ l'espérance associée.

Fixons une fonction $f : F \to \mathbb{R}$ mesurable telle que $0 < \inf f \leq f \leq 1$ et posons pour tout $(t,x) \in \mathbb{R}_+ \times F$

$$w(t,x) := \mathbb{E}_{\delta_x}\big[e^{\langle Z_t, \log f\rangle}\big] = \mathbb{E}_{\delta_x}\Bigg[\prod_{i \in I(t)} f(\xi_i(t))\Bigg]. \tag{1.8}$$

On a alors la propriété suivante.

PROPOSITION 1.5. *La fonction w est solution de l'équation intégrale*

$$w(t,x) - \rho\, \mathrm{E}_x\bigg[\int_0^t \Big\{\Phi\Big(w\big(t-s,\xi_s\big)\Big) - w\big(t-s,\xi_s\big)\Big\}ds\bigg] = \mathrm{E}_x\big[f(\xi_t)\big]. \tag{1.9}$$

On comparera cette équation à celle donnée dans (1.2).

Éléments de preuve. La propriété de branchement nous permet d'écrire que si $n \in \mathbb{N}$,

$$\mathbb{E}_{n\delta_x}\Big[e^{\langle Z_t, \log f\rangle}\Big] = w(t,x)^n \tag{1.10}$$

(utiliser la troisième expression dans (1.8) pour s'en convaincre). En introduisant le premier temps de branchement et en découpant l'intégrale définissant $w(t,x)$ selon que ce temps est plus grand ou plus petit que t, on obtient

$$w(t,x) = e^{-\rho t}\mathrm{E}_x\big[f(\xi_t)\big] + \rho\, \mathrm{E}_x\bigg[\int_0^t ds\, e^{-\rho s}\Phi\big(w(t-s,\xi_s)\big)\bigg],$$

où le premier terme correspond au cas où l'unique particule de Z_0 ne branche pas avant l'instant t (mais se déplace selon la loi du processus ξ), et le second terme décrit le cas où cette particule initiale évolue dans F selon la loi de ξ jusqu'au moment s auquel elle branche (distribué selon une variable exponentielle de paramètre ρ). Puis les K descendants créés partent tous de la position ξ_s et évoluent suivant des copies indépendantes de Z pendant le temps $t-s$ restant. Observons que l'on a utilisé (1.10) pour introduire la fonction génératrice de K dans cette expression. En manipulant le second terme de l'égalité ci-dessus, on arrive rapidement à l'égalité (1.9). Une preuve complète est donnée dans [**LG99**], p.29. ☐

Comme souvent dans les processus de branchement, un système fini de particules n'est pas facile à décrire et faire tendre la taille de la population vers l'infini conduit à un objet plus maniable mais qui retient cependant certaines caractéristiques essentielles des populations finies. C'est donc la ligne de conduite que nous allons suivre à présent et qui nous amènera à la définition du super-mouvement brownien. Supposons ρ et Φ fixés, et en outre que Φ satisfait $\Phi'(1) = 1$ et $\Phi''(1) < \infty$ (on se restreint donc à des branchements critiques, de variance σ^2 du nombre de descendants finie). On revient au cas où chaque particule se déplace selon la loi d'un mouvement brownien dans $\mathbb{R}^d$.

La manière adéquate de faire tendre le nombre de particules vers l'infini est fortement suggérée par la renormalisation opérée à la fin du paragraphe précédent pour obtenir une diffusion de Feller. Procédons donc de la même manière et définissons pour tout $n \in \mathbb{N}$ le processus aléatoire Z^n à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$ par : pour toute fonction f mesurable bornée et tout $t \geq 0$,

$$\langle Z_t^n, f\rangle := \frac{1}{n}\Big\langle Z_{nt}^{[n]}, f\Big(\frac{\cdot}{\sqrt{n}}\Big)\Big\rangle, \tag{1.11}$$

où $(Z^{[n]})_{n\in\mathbb{N}}$ est une suite de mouvements browniens branchants de mêmes paramètres ρ et Φ. De manière équivalente, on peut écrire

$$Z_t^n = \frac{1}{n} \sum_{i\in I^{[n]}(nt)} \delta_{\xi_i(nt)/\sqrt{n}}.$$

Remarquons tout d'abord qu'une renormalisation des positions spatiales des particules est également nécessaire. En effet, celles-ci se déplacent n fois plus vite si on accélère le temps par un facteur n et parcourent donc une distance d'ordre $\sqrt{n}$ en ce qui est à présent une unité de temps. La propriété d'invariance par changement d'échelle du mouvement brownien assure alors que chaque particule de Z^n se déplace toujours suivant la loi d'un mouvement brownien.

Supposons que $(Z_0^n)_{n\in\mathbb{N}}$ converge en loi vers une mesure aléatoire $X_{(0)}$ à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$. On a alors le résultat suivant.

DÉFINITION ET THÉORÈME 1.6. *La suite $(Z^n)_{n\in\mathbb{N}}$ converge (au sens de la convergence des marginales de dimension finie) vers un processus markovien X à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$, dont la valeur initiale est distribuée selon la loi de $X_{(0)}$ et dont les noyaux de transition $Q_t(\mu, d\nu)$ sont caractérisés par la propriété suivante : pour tout $\mu \in \mathcal{M}_f(\mathbb{R}^d)$ et toute fonction $f \in \mathcal{B}_b^+(\mathbb{R}^d)$,*

$$\int Q_t(\mu, d\nu)\, e^{-\langle \nu, f\rangle} = e^{-\langle \mu, u(t,\cdot)\rangle}, \tag{1.12}$$

où la fonction $u(t,x)$ est l'unique solution positive de l'équation intégrale

$$u(t,x) + \mathrm{E}_x\left[\int_0^t \psi\big(u(t-s,\xi_s)\big)ds\right] = \mathrm{E}_x\big[f(\xi_t)\big]. \tag{1.13}$$

On a noté ici $\psi(u) = \frac{\rho\sigma^2}{2}u^2$. Le processus X est appelé processus de Dawson-Watanabe *ou* super-mouvement brownien.

En outre, X satisfait la propriété de branchement

$$Q_t(\mu, \cdot) * Q_t(\mu', \cdot) = Q_t(\mu + \mu', \cdot).$$

On comparera cette fois (1.13) à l'équation (1.4).

Plus généralement, si ξ est un processus markovien sur un espace polonais F et ψ est une fonction de la forme (1.5), il existe un unique (en loi) processus markovien à valeurs dans $\mathcal{M}_f(F)$ dont les noyaux de transition satisfont (1.12) et (1.13). On l'appelle un (ξ, ψ)-superprocessus. Dans le cas précédent où $\psi(u) = cu^2$, on parlera de *branchement quadratique*.

REMARQUE 1.7. *En choisissant des fonctions f constantes, on obtient que la masse totale $\langle X, 1\rangle$ d'un (ξ,ψ)-superprocessus est un CSBP de mécanisme de branchement ψ.*

Éléments de preuve du Théorème 1.6. Laissons de côté la preuve de l'existence et l'unicité de la solution positive de (1.13) et du fait que ces équations intégrales caractérisent bien les noyaux de transition d'un processus markovien à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$. On la trouvera dans [**LG99**], p.30, pour une suite $(Z^n)_{n\in\mathbb{N}}$ légèrement différente de celle que nous avons utilisée.

Donnons simplement quelques clés pour comprendre d'où vient la convergence des mouvements browniens branchants renormalisés vers le super-mouvement brownien. On se limitera à la convergence des marginales unidimensionnelles, dont (1.9) nous fournit une caractérisation. Fixons donc une fonction $f \in \mathcal{B}_b^+(\mathbb{R}^d)$ et posons pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}^d$

$$g_n(x) := \exp\big(-n^{-1}f(x/\sqrt{n})\big).$$

Si on définit également pour tout n la fonction $w_n(t,x)$ correspond à $Z^{[n]}$ et g_n comme dans (1.8), on a alors d'une part

$$w_n(nt,\sqrt{n}x) = \mathbb{E}_{\delta_{\sqrt{n}x}}\Big[\exp\Big(-n^{-1}\big\langle Z^{[n]}_{nt}, f(\cdot/\sqrt{n})\big\rangle\Big)\Big] = \mathbb{E}_{\delta_x}\Big[e^{-\langle Z^n_t, f\rangle}\Big],$$

et d'autre part d'après (1.9),

$$\begin{aligned}
w_n(nt,\sqrt{n}x) &= \rho\,\mathrm{E}_{\sqrt{n}x}\left[\int_0^{nt}\Big\{\Phi\left(w_n\big(nt-s,\xi_s\big)\right) - w_n\big(nt-s,\xi_s\big)\Big\}ds\right] \\
&\quad + \mathrm{E}_{\sqrt{n}x}\Big[e^{-\frac{1}{n}f(\xi_{nt}/\sqrt{n})}\Big] \\
&= \rho n\,\mathrm{E}_x\left[\int_0^{t}\Big\{\Phi\left(w_n\big(n(t-s),\sqrt{n}\xi_s\big)\right) - w_n\big(n(t-s),\sqrt{n}\xi_s\big)\Big\}ds\right] \\
&\quad + \mathrm{E}_x\Big[e^{-\frac{1}{n}f(\xi_t)}\Big],
\end{aligned}$$

où la dernière égalité découle du changement de variable temporelle $s' = s/n$ et de la propriété d'invariance par changement d'échelle du mouvement brownien. Si on pose à présent

$$u_n(t,x) := n\big(1 - w_n(nt,\sqrt{n}x)\big),$$

on obtient que u_n satisfait

$$\begin{aligned}
u_n(t,x) &= n\left(1 - \mathrm{E}_x\Big[e^{-\frac{1}{n}f(\xi_t)}\Big]\right) \\
&\quad -\rho n^2\,\mathrm{E}_x\left[\int_0^t\left\{\Phi\left(1 - \frac{u_n(t-s,\xi_s)}{n}\right) - \left(1 - \frac{u_n(t-s,\xi_s)}{n}\right)\right\}ds\right],
\end{aligned}\tag{1.14}$$

et en utilisant un développement de Taylor de Φ au voisinage de 1 et les égalités $\Phi'(1) = 1$ et $\Phi''(1) = \sigma^2$, on obtient l'approximation

$$u_n(t,x) \approx \mathrm{E}_x\big[f(\xi_t)\big] - \frac{\rho\sigma^2}{2}\,\mathrm{E}_x\left[\int_0^t u_n(t-s,\xi_s)^2 ds\right]$$

lorsque n est grand. Pour rendre cette approximation rigoureuse, remarquons d'abord que la positivité de $\Phi - \mathrm{Id}$ et l'égalité (1.14) garantissent que u_n est bornée par la norme sup de f pour tout n, et donc on a bien

$$\Phi\left(1 - \frac{u_n(t-s,\xi_s)}{n}\right) - \left(1 - \frac{u_n(t-s,\xi_s)}{n}\right) = \frac{\sigma^2}{2n^2}\,u_n(t-s,\xi_s)^2 + o\Big(\frac{1}{n^2}\Big)$$

lorsque $n \to \infty$. En utilisant les équations intégrales satisfaites par u et u_n, ainsi que le fait que ψ est lipschitzienne sur tout intervalle compact et que f est bornée, on peut ainsi écrire pour n assez grand, $t \in [0,T]$ et $x \in \mathbb{R}^d$

$$\big|u_n(t,x) - u(t,x)\big| \le C\int_0^t \sup_{y\in\mathbb{R}^d}\big|u_n(s,y) - u(s,y)\big|\,dy + \varepsilon(n,T,f),\tag{1.15}$$

où $\varepsilon(n,T,f) \to 0$ lorsque $n \to \infty$. Le lemme de Gronwall assure alors que $(u_n)_{n\in\mathbb{N}}$ converge uniformément vers u sur tout compact de la forme $[0,T] \times \mathbb{R}^d$. Puisque la convergence des transformées de Laplace d'une suite de variables aléatoires réelles implique la convergence en loi de ces variables, on en déduit que si $t \geq 0$ est fixé, alors pour toute fonction $f \in \mathcal{B}_b^+(\mathbb{R}^d)$, $(\langle Z^n_t, f\rangle)_{n\in\mathbb{N}}$ converge en loi vers $\langle X_t, f\rangle$. L'ensemble des fonctions de la forme $\mu \mapsto \langle \mu, f\rangle$ pour $f \in \mathcal{B}_b^+(\mathbb{R}^d)$ *détermine la convergence* pour des variables aléatoires à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$ et donc ceci est suffisant pour conclure que $(Z^n_t)_{n\in\mathbb{N}}$ converge en loi vers X_t.

Notons pour finir que l'uniformité de la convergence de u_n nous permet d'obtenir la convergence des marginales de dimension finie et que la propriété de branchement de X se vérifie facilement à partir de (1.12). $\square$

On conserve la notation $\mathbb{P}_\mu$ pour la mesure de probabilité sous laquelle $X_0 = \mu$ p.s.

REMARQUE 1.8. *On peut généraliser (1.13) à une équation caractérisant les marginales de dimension finie du super-mouvement brownien. En effet, notons* $\mathrm{P}_{t,x}$ *la mesure de probabilité sous laquelle le mouvement brownien* ξ *part de* x *au temps* t*, avec la convention que* $\mathrm{E}_{t,x}[f(\xi_s)] = 0$ *si* $s < t$*. Alors, si* $0 \leq t_1 < \ldots < t_k$ *et* $f_1, \ldots, f_k \in \mathcal{B}_b^+(\mathbb{R}^d)$*, on a*

$$\mathbb{E}_\mu\left[\exp\left\{-\sum_{i=1}^k \langle X_{t_i}, f_i\rangle\right\}\right] = e^{-\langle \mu, w(0,\cdot)\rangle},$$

où la fonction $w(t,x)$ *est l'unique solution positive de l'équation intégrale*

$$w(t,x) + \mathrm{E}_{t,x}\left[\int_t^\infty \psi\big(w(s,\xi_s)\big)\, ds\right] = \mathrm{E}_{t,x}\left[\sum_{i=1}^k f_i\big(\xi_{t_i}\big)\right].$$

Une preuve de ce résultat est donnée dans [**LG99**]*, Proposition II.7.*

Par les mêmes arguments que pour les CSBP, le super-mouvement brownien vérifie la propriété de Markov forte et possède une modification à trajectoires càdlàg (pour la topologie de la convergence faible dont on munit $\mathcal{M}_f(\mathbb{R}^d)$). En fait, il existe même une modification à trajectoires continues (cf. [**Eth00**], Théorème 2.13). Citons également quelques propriétés supplémentaires, qui donnent une idée de ce à quoi ressemble X à un temps t fixé :

- en dimension 1, X_t a une densité par rapport à la mesure de Lebesgue que l'on notera de manière abusive $X_t(x)$. Si le mécanisme de branchement de X est $\psi(u) = \beta u^2$, cette densité est une solution faible de l'équation différentielle stochastique
$$dX_t = \frac{1}{2}\Delta X_t dt + \sqrt{2\beta X_t}\, dB_t;$$
- en dimension supérieure, X_t est singulière par rapport à la mesure de Lebesgue et son support est de dimension de Hausdorff 2 en toute dimension $d \geq 2$.

On consultera le chapitre 2 de [**Eth00**] pour une preuve et des compléments.

Enfin, notons $\{T_t,\ t \geq 0\}$ le semi-groupe de transition du mouvement brownien, i.e., $T_t f(x) = \mathrm{E}_x[f(\xi_t)]$. On pourra trouver la preuve du résultat suivant dans le chapitre II.4 de [**LG99**] :

PROPOSITION 1.9. *Soit* X *un super-mouvement brownien de mécanisme de branchement* ψ *de la forme (1.5). Alors pour toutes fonctions* $f, g \in \mathcal{B}_b^+(\mathbb{R}^d)$ *et* $t, t' \geq 0$*,*

a)

$$\mathbb{E}_\mu\big[\langle X_t, f\rangle\big] = e^{-\alpha t}\langle \mu, T_t f\rangle.$$

b) Si $\psi(u) = \beta u^2$*,*

$$\mathbb{E}_\mu\big[\langle X_t, f\rangle\langle X_{t'}, g\rangle\big] = \langle \mu, T_t f\rangle\langle \mu, T_{t'} g\rangle + 2\beta \int_0^{t\wedge t'} \big\langle \mu, T_s\big((T_{t-s}f)(T_{t'-s}g)\big)\big\rangle\, ds,$$

et en particulier

$$\mathbb{E}_\mu\big[\langle X_t, f\rangle^2\big] = \langle \mu, T_t f\rangle^2 + 2\beta \int_0^t \big\langle \mu, T_s\big((T_{t-s}f)^2\big)\big\rangle\, ds.$$

Le moment d'ordre 1 de $\langle X_t, f\rangle$ est donc donné par le comportement d'une seule particule brownienne, alors que celui d'ordre 2 s'exprime en fonction de deux particules « échantillonnées dans la population au temps t », qui proviennent de deux particules distinctes au temps 0 se déplaçant indépendamment (premier terme), ou au contraire qui sont nées à un temps $s \in [0,t]$ de la même particule brownienne et évoluent ensuite indépendamment pendant le temps $t-s$ restant (second terme). En fait, si le mécanisme de branchement de X est quadratique, on peut exprimer le moment d'ordre k de $\langle X_t, f\rangle$ au moyen d'une somme faisant intervenir toutes les généalogies possibles de k particules (cf. [**Eth00**], chapitre 2.2).

REMARQUE 1.10. (Superprocessus historique) *On peut garder une trace de la généalogie et des trajectoires passées de toutes les « particules » en vie à l'instant* t *en prenant comme espace* F *l'ensemble* $\mathcal{W}$ *des chemins (càdlàg) finis à valeurs dans* $\mathbb{R}^d$. *On remplace alors le processus* ξ *par* $\overline{\xi}$ *défini par*

$$\overline{\xi}_t := \big(\xi_s,\ s \in [0,t]\big),$$

qui est également un processus markovien cette fois à valeurs dans $\mathcal{W}$. *On appelle le* $(\overline{\xi},\psi)$*-superprocessus* $\mathcal{H}$ *ainsi obtenu* superprocessus historique. *Notons que l'on peut exprimer le superprocessus initial* X *en fonction de* $\mathcal{H}$ *en ne considérant que le point final de chaque trajectoire arrêtée. En effet, définissons* $\tilde{X}$ *par : pour toute fonction* f *mesurable bornée et tout* $t \geq 0$,

$$\langle \tilde{X}_t, f\rangle := \int_{\mathcal{W}} f(w_t)\, \mathcal{H}_t(dw).$$

Alors $\tilde{X} \overset{(d)}{=} X$, *i.e.,* $\tilde{X}$ *est un* (ξ,ψ)*-superprocessus. Une bonne référence pour l'étude des superprocessus historiques est* [**DP91**].

La littérature sur les superprocessus est incroyablement vaste et les super-mouvements browniens de branchement quadratique en particulier ont fait l'objet d'études très détaillées. On renvoie à [**Daw93**], [**Eth00**], [**LG99**] et [**Per02**] pour une revue d'un bon nombre de ces résultats.

1.2.2. *Caractérisation via un problème de martingales.* Une autre caractérisation des processus de branchement spatiaux du paragraphe précédent peut s'avérer utile : celle en termes de générateur et de problème de martingales développée par Stroock et Varadhan (cf. [**SV79**]). En particulier, on pourra se convaincre dans quelques lignes qu'elle est très bien adaptée aux problèmes de convergence de processus aléatoires.

De manière informelle, le générateur d'un processus markovien $\{Y_t,\ t \geq 0\}$ à valeurs dans un espace F est l'opérateur linéaire $\mathcal{L}$ sur les fonctions $f : F \to \mathbb{R}$ donné par

$$\mathcal{L}f(y) := \lim_{s\to 0} \frac{\mathbb{E}_y[f(Y_s)] - f(y)}{s} = \frac{d}{ds}\mathbb{E}_y[f(Y_s)]\Big|_{s=0}, \qquad y \in F. \tag{1.16}$$

Le domaine $\mathcal{D}(\mathcal{L})$ de $\mathcal{L}$ (i.e., son ensemble de définition) est l'ensemble des fonctions f telles que (1.16) est fini pour tout y. En utilisant la propriété de Markov de Y, on obtient également que pour tout $t \geq 0$,

$$\frac{d}{ds}\mathbb{E}_y[f(Y_s)]\Big|_{s=t} = \mathbb{E}_y[\mathcal{L}f(Y_t)],$$

d'où en intégrant entre les instants t et $t+u$ et en appliquant la propriété de Markov au temps t :

$$\mathbb{E}_y\big[f(Y_{t+u}) - f(Y_t)\big|\mathcal{F}_t\big] = \mathbb{E}_y\bigg[\int_t^{t+u} \mathcal{L}f(Y_s)\, ds \,\bigg|\, \mathcal{F}_t\bigg],$$

où $\mathcal{F} = \{\mathcal{F}_t,\ t \geq 0\}$ est la filtration canonique associée à Y. Modulo des conditions d'intégrabilité, on en déduit que

$$f(Y_t) - f(Y_0) - \int_0^t \mathcal{L}f(Y_s)\, ds$$

est une martingale pour la filtration $\mathcal{F}$. Ceci nous conduit à la définition suivante.

DÉFINITION 1.11. *On dit que le processus Y est* solution du problème de martingales associé à $\mathcal{L}$ *si pour tout* $f \in \mathcal{D}(\mathcal{L})$,

$$f(Y_t) - f(Y_0) - \int_0^t \mathcal{L}f(Y_s)\, ds$$

est une martingale (pour la filtration canonique associée à Y).

Il n'est pas difficile de voir qu'un problème de martingales ne fournit d'informations que sur les marginales de dimension finie de Y. On dira donc qu'il est *bien posé* s'il admet au moins une solution et que toutes ses solutions ont les mêmes marginales fini-dimensionnelles (en loi). En fait, le Théorème 4.4.2 de [**EK86**] nous dit que si F est un espace séparable, il suffit de vérifier l'égalité en loi des marginales unidimensionnelles et que, si elle est satisfaite, toute solution est alors markovienne. On renvoie aux chapitres 4.3 et 4.4 de [**EK86**] pour des compléments, notamment sur les propriétés des solutions de problèmes de martingales bien posés.

Pour commencer, écrivons un problème de martingales correspondant à un mouvement brownien branchant.

THÉORÈME 1.12. *Le mouvement brownien branchant Z de paramètres ρ et Φ est l'unique solution du problème de martingales : pour tout $f \in \mathcal{D}(\Delta)$ telle que* $0 < \inf f \leq f \leq 1$,

$$e^{\langle Z_t, \log f\rangle} - e^{\langle Z_0, \log f\rangle} - \int_0^t \left\langle Z_s\,,\ \frac{\frac{1}{2}\Delta f + \rho\big(\Phi(f) - f\big)}{f}\right\rangle e^{\langle Z_s, \log f\rangle}\, ds \tag{1.17}$$

est une martingale.

Éléments de preuve. On ne donne ici que les idées permettant d'obtenir le générateur de Z et on renvoie aux commentaires à la fin du chapitre 1.2 de [**Eth00**] pour la question de l'unicité. Pour commencer, rappelons-nous qu'en utilisant l'écriture (1.6), on a

$$e^{\langle Z_t, \log f\rangle} = \prod_{i \in I(t)} f(\xi_i(t)).$$

Définissons $w(t,x)$ comme dans (1.8). Par des arguments analytiques standards, la Proposition 1.5 assure que w est solution du système

$$\begin{cases} \frac{\partial w}{\partial t} = \frac{1}{2}\Delta w + \rho\big(\Phi(w) - w\big), \\ w(0,x) = f(x). \end{cases}$$

En utilisant la propriété de Markov au temps $t \geq 0$ et un théorème de dérivation sous le signe intégral, on arrive à

$$\begin{aligned} \frac{d}{ds}\mathbb{E}\big[e^{\langle Z_s, \log f\rangle}\big]\Big|_{s=t} &= \mathbb{E}\Bigg[\sum_{i \in I(t)} \frac{1}{w(0,\xi_i(t))}\frac{\partial w}{\partial t}(0,\xi_i(t))\, e^{\langle Z_t, \log f\rangle}\Bigg] \\ &= \mathbb{E}\Bigg[\left\langle Z_t\,,\ \frac{\frac{1}{2}\Delta f + \rho\big(\Phi(f) - f\big)}{f}\right\rangle e^{\langle Z_t, \log f\rangle}\Bigg]. \end{aligned} \tag{1.18}$$

On achève alors le raisonnement de la même manière que dans le paragraphe précédant la Définition 1.11, en remarquant que l'intégrabilité de notre martingale potentielle découle des hypothèses faites sur f. □

Par-delà le calcul, on notera que la division par $f(\xi_i(t))$ dans (1.18) permet de retirer cette valeur de $e^{\langle Z_t, \log f\rangle}$ et que le numérateur correspondant donne ensuite l'incrément infinitésimal de $f(\xi_i(t))$: soit la particule i se déplace simplement sur l'intervalle de temps $[t, t + dt]$ (d'où le générateur du mouvement brownien), soit elle branche et produit un nombre de particules donné par Φ (on peut alors négliger le fait qu'elle se soit aussi déplacée). On somme ensuite les contributions de chaque particule en vie à l'instant t à l'évolution instantanée de $\Pi_i f(\xi_i(t))$.

Voyons ensuite ce qu'il arrive à la suite $(Z^n)_{n\in\mathbb{N}}$ définie par (1.11). Comme dans le paragraphe précédent, on suppose que $\Phi'(1) = 1$ et $\sigma^2 = \Phi''(1) < \infty$. Posons cette fois $g_n = 1 - \frac{1}{n} f\big(\frac{\cdot}{\sqrt{n}}\big)$ pour une fonction $f \in \mathcal{D}(\Delta)$ positive, bornée et telle que Δf est aussi bornée. Le problème de martingales (1.17) écrit au temps nt avec la fonction g_n nous donne que

$$\exp\Big\{ \Big\langle nZ_t^n, \log\Big(1 - \frac{f}{n}\Big)\Big\rangle\Big\} - \exp\Big\{\Big\langle nZ_0^n, \log\Big(1 - \frac{f}{n}\Big)\Big\rangle\Big\}$$
$$- \int_0^{nt} \Big\langle nZ^n_{s/n}\,,\ \frac{-\frac{1}{2n^2}\Delta f + \rho\big(\Phi(1 - \frac{f}{n}) - 1 + \frac{f}{n}\big)}{1 - \frac{f}{n}}\Big\rangle \exp\Big\{\Big\langle nZ^n_{s/n}, \log\Big(1 - \frac{f}{n}\Big)\Big\rangle\Big\}\, ds$$

est une martingale (pour la filtration canonique de Z^n, i.e., $\{\mathcal{F}^{[n]}_{nt},\ t \geq 0\}$). On obtient alors que l'expression ci-dessus est égale à

$$e^{-\langle Z_t^n, f\rangle} - e^{-\langle Z_0^n, f\rangle} - \int_0^t \Big\langle Z_s^n, -\frac{1}{2}\Delta f + \frac{\rho\sigma^2}{2}\, f^2\Big\rangle\, e^{-\langle Z_s^n, f\rangle}\, ds \tag{1.19}$$

à un terme d'erreur de la forme $t.\varepsilon(n, \|f\|, \|\Delta f\|)$ près, où $\varepsilon(n, \|f\|, \|\Delta f\|) \to 0$ quand $n \to \infty$ (on a noté ici $\|g\|$ la norme sup de g). Ce calcul ne garantit en rien que toute limite de la suite $(Z^n)_{n\in\mathbb{N}}$ satisfasse le problème de martingales (1.19), ni d'ailleurs que celui-ci soit bien posé. Cependant, on peut montrer que ces deux énoncés sont vrais (cf. paragraphe 1.2.3 pour le premier et chapitre 1.6 de [**Eth00**] pour le second), ce qui nous conduit à la définition suivante :

DÉFINITION 1.13. *On dit que le processus X à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$ satisfait le problème de martingales de Dawson-Watanabe s'il existe $\beta > 0$ tel que pour toute fonction $f \in \mathcal{D}(\Delta)$,*

$$e^{-\langle X_t, f\rangle} - e^{-\langle X_0, f\rangle} - \int_0^t \Big\langle X_s, -\frac{1}{2}\Delta f + \beta f^2\Big\rangle\, e^{-\langle X_s, f\rangle}\, ds$$

est une martingale.

Comme on s'y attend, le super-mouvement brownien de mécanisme de branchement $\psi(u) = \beta u^2$ satisfait le problème de martingales de Dawson-Watanabe de paramètre β et puisque ce dernier est bien posé, il en caractérise les marginales de dimension finie.

1.2.3. *Convergence de processus à valeurs mesures.* Il est temps de dire un mot des différents espaces dans lesquels nous travaillons et des arguments nécessaires pour montrer des résultats de convergence dans ces espaces de dimension infinie. Une référence très complète sur ce sujet est [**EK86**] (le chapitre 3 en particulier) ; on pourra également consulter [**Bil68**] pour une revue plus synthétique.

On notera $C_b(F)$ l'ensemble des fonctions réelles continues et bornées définies sur l'espace F, et $\|f\|$ la norme sup de $f \in C_b(F)$. Commençons par décrire les espaces qui nous intéressent.

1. Les processus que l'on étudie sont à valeurs dans l'ensemble $\mathcal{M}_f(\mathbb{R}^d)$ des mesures finies sur $\mathbb{R}^d$, que l'on munit de la topologie de la convergence faible (notée $\Rightarrow$). Cette topologie fait de $\mathcal{M}_f(\mathbb{R}^d)$ un espace polonais (métrisable, séparable et complet).
2. On ne considère que leurs modifications à trajectoires continues à droite avec limite à gauche, dont la loi est donc une mesure de probabilité sur l'espace $D_{\mathcal{M}_f(\mathbb{R}^d)}[0, \infty)$ des fonctions càdlàg de $[0, \infty)$ dans $\mathcal{M}_f(\mathbb{R}^d)$. On notera $\mathcal{M}_1(D_{\mathcal{M}_f(\mathbb{R}^d)}[0, \infty))$ l'ensemble des mesures de probabilité sur cet espace, et plus généralement $\mathcal{M}_1(F)$ les mesures de probabilité sur un espace mesurable F. On munit $D_{\mathcal{M}_f(\mathbb{R}^d)}[0, \infty)$ d'une distance particulière (dont on trouvera la définition dans [**EK86**], p.117) et la topologie associée est appelée

topologie de Skorokhod. Le fait important ici est que d'après le Théorème 3.5.6 de [**EK86**] et le point 1., on obtient alors également un espace polonais.

3. On munit $\mathcal{M}_1(D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty))$ d'une distance pour laquelle la convergence d'une suite d'éléments de $\mathcal{M}_1(D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty))$ est équivalente à leur convergence faible. L'intérêt de la topologie associée à cette distance est qu'on sait en caractériser les ensembles compacts.

Tous ces résultats sont tirés du chapitre 3 de [**EK86**].

La preuve de la convergence en loi de processus markoviens à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$ se divise en deux étapes : on montre d'abord que la suite est tendue, puis que toutes les sous-suites convergentes admettent nécessairement la même limite.

Tension

DÉFINITION 1.14. *Soit (S,d) un espace métrique. Une famille $M \subset \mathcal{M}_1(S)$ est dite* tendue *si pour tout $\varepsilon > 0$, il existe un compact $K \subset S$ tel que pour tout $\mu \in M$, $\mu(K) > 1-\varepsilon$. Par extension, une famille de variables aléatoires à valeurs dans S est dite tendue si leurs lois le sont.*

Pour nous $S = D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$, qui est comme on l'a vu un espace polonais. Par conséquent, d'après le Théorème 3.2.2 de [**EK86**] la tension est équivalente à la compacité relative de la famille de mesures de probabilité considérée. Mais grâce au choix judicieux de la distance dont on a muni S (voir le point 3. ci-dessus), on dispose d'un critère de compacité relative. Avant de l'énoncer, introduisons deux notations. Supposons que F soit un espace polonais muni d'une distance d. On définit pour tout $X \in D_F[0,\infty)$, $\delta > 0$ et $T > 0$ le *module de continuité* $w'(X,\delta,T)$ de X sur l'intervalle $[0,T]$ et de pas δ par

$$w'(X,\delta,T) := \inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} d\big(X_s,X_t\big), \tag{1.20}$$

où la collection finie $\{t_i\}$ de temps est de la forme $0 = t_0 < t_1 < \ldots < t_m = T$ pour un entier $m > 0$ et satisfait $\min_{1\le i\le m}(t_i - t_{i-1}) > \delta$. De manière heuristique, on veut pouvoir placer les sauts (ou les variations) de X trop importants aux extrémités de la subdivision pour garder le supremum sur s,t de (1.20) petit, et ceci n'est possible que si ces sauts ne s'accumulent pas autour d'un ou plusieurs points. Enfin, on note A^ε le ε-voisinage de $A \subset F$, i.e., l'ensemble des points à distance plus petite que ε d'un élément de A.

THÉORÈME 1.15. (Corollaire 3.7.4 de [**EK86**]) *Soit $(X^n)_{n\in\mathbb{N}}$ une suite de processus à trajectoires càdlàg dans F. Alors $(X^n)_{n\in\mathbb{N}}$ est relativement compacte ssi les deux conditions suivantes sont satisfaites :*

1. *Pour tout $\eta > 0$ et $t \ge 0$ (rationnel), il existe un compact $\Gamma_{\eta,t} \subset F$ tel que*
$$\liminf_{n\to\infty} \mathbb{P}\big[X_t^n \in \Gamma_{\eta,t}^\eta\big] \ge 1-\eta.$$
2. *Pour tout $\eta > 0$ et $T > 0$, il existe $\delta > 0$ tel que*
$$\limsup_{n\to\infty} \mathbb{P}\big[w'(X^n,\delta,T) \ge \eta\big] \le \eta. \tag{1.21}$$

L'espace d'intérêt pour nous est $F = \mathcal{M}_f(\mathbb{R}^d)$. Cependant, comme on peut s'en douter la distance dont on munit $\mathcal{M}_f(\mathbb{R}^d)$ n'est pas particulièrement simple à manipuler. Heureusement, le théorème suivant nous permet de nous ramener à des processus à valeurs réelles.

THÉORÈME 1.16. (Théorème 3.9.1 de [**EK86**]) *Soient F un espace polonais et $(X^n)_{n\ge1}$ une suite de processus à trajectoires càdlàg dans F. Supposons que la* condition de confinement dans un compact *est satisfaite, i.e., que pour tout* $\eta > 0$ *et* $T > 0$*, il existe un compact* $\Gamma_{\eta,T} \subset F$ *tel que*

$$\inf_{n\in\mathbb{N}} \mathbb{P}\big[X_t^n \in \Gamma_{\eta,T} \text{ pour tout } 0 \le t \le T\big] \ge 1 - \eta.$$

Soit H *un ensemble dense dans* $C_b(F)$ *pour la topologie de la convergence uniforme sur les compacts. Alors* $(X^n)_{n\in\mathbb{N}}$ *est relativement compacte ssi* $(f(X^n))_{n\in\mathbb{N}}$ *est relativement compacte pour tout* $f \in H$*.*

On peut ensuite appliquer le Théorème 1.15 à chaque suite $(f(X^n))_{n\in N}$, dont chacun des termes a des trajectoires càdlàg à valeurs réelles. La suite de la démonstration dépend du cadre de travail : soit on montre (1.21) directement comme on le fera dans le chapitre 2, soit on utilise des critères appropriés comme dans la preuve du Lemme 3.10 du chapitre 3.

Unicité des valeurs d'adhérences

La deuxième étape consiste à caractériser les valeurs d'adhérence pour montrer qu'il en existe au plus une. Pour ce faire, on utilise le fait que la tribu sur $D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$ engendrée par les ouverts de la topologie de Skorokhod est égale à celle engendrée par la famille des projections $\pi_t : X \mapsto X_t$ (cf. Proposition 3.7.1 de [**EK86**]). Autrement dit, la loi d'un processus à trajectoires dans $D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$ est caractérisée par l'ensemble des lois de ses marginales de dimension finie. Mais c'est précisément une caractérisation de ces lois que nous offrent les équations d'évolution ou les problèmes de martingales des paragraphes 1.2.1 et 1.2.2. On peut donc utiliser les résultats suivants pour obtenir une version plus forte du Théorème 1.6, à savoir que la loi sur $D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$ de Z^n converge vers celle de X lorsque $n \to \infty$. Les deux théorèmes ci-dessous sont des cas (très) particuliers des Théorèmes 3.7.8 et 4.8.10 de [**EK86**] respectivement.

THÉORÈME 1.17. *Soit* $(X^n)_{n\in\mathbb{N}}$ *et* X *des processus à trajectoires dans* $D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$*. Si* $(X^n)_{n\in\mathbb{N}}$ *est relativement compacte et si pour tout* $0 \le t_1 < t_2 < \ldots < t_k$ *on a*

$$(X_{t_1}^n, \ldots, X_{t_k}^n) \overset{(d)}{\to} (X_{t_1}, \ldots, X_{t_k})$$

lorsque $n \to \infty$*, alors* $X^n \overset{(d)}{\to} X$*.*

THÉORÈME 1.18. *Soit* $\mathcal{L}$ *un opérateur linéaire sur* $C_b(\mathcal{M}_f(\mathbb{R}^d))$ *et supposons qu'il existe au plus une solution au problème de martingales associé à* $\mathcal{L}$*. Soit* $(X^n)_{n\in\mathbb{N}}$ *une suite de processus à trajectoires dans* $D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$ *. Si* $(X^n)_{n\in\mathbb{N}}$ *est relativement compacte et si pour tout* $f \in \mathcal{D}(\mathcal{L})$*,*

$$\lim_{n\to\infty} \mathbb{E}\left[\left(f(X_{t+s}^n) - f(X_t^n) - \int_t^{t+s} \mathcal{L}f(X_u^n)\,du\right) \prod_{i=1}^k h_i(X_{t_i}^n)\right] = 0$$

pour tous $k \ge 0$*,* $0 \le t_1 < t_2 < \ldots < t_k \le t < t+s$ *et* $h_1, \ldots, h_k \in C_b(\mathcal{M}_f(\mathbb{R}^d))$*, alors il existe une solution* X *au problème de martingales associé à* $\mathcal{L}$ *et* $X^n \overset{(d)}{\to} X$*.*

Il existe de nombreux raffinements de ces résultats, mais une fois de plus on renvoie aux chapitres 3 et 4 de [**EK86**] pour une exposition très complète. Le Théorème 1.17 nous servira dans le chapitre 2 et le Théorème 1.18 pour le résultat de convergence du chapitre 3.

1.3. Mouvement brownien (branchant) parmi des obstacles aléatoires. Jusqu'alors l'évolution de nos populations ne dépendait pas de leur « environnement », i.e., nous avons supposé que les lois de reproduction et de déplacement ne dépendaient ni du temps, ni de la position spatiale des individus. L'objectif de la suite de ce paragraphe et des chapitres 2 et 3 de cette thèse est d'introduire une variation aléatoire des conditions environnementales d'un type particulier, et de comprendre quels en sont les effets sur les paramètres dictant l'évolution des processus de branchement spatiaux correspondant.

Trois types de processus de branchement en milieu aléatoire sont particulièrement étudiés :

Processus de Galton-Watson (non-spatiaux): On suppose que la loi de reproduction est la même pour tous les individus, mais est tirée pour chaque génération de manière i.i.d. parmi un ensemble de lois prédéfini. D'une génération à l'autre, les individus ne se reproduisent donc pas selon le même mécanisme de branchement, mais on conserve malgré tout une certaine cohérence du fait que ces lois ont elles-mêmes des distributions i.i.d. Ce modèle est bien évidemment généralisable aux processus spatiaux.

Processus de Galton-Watson spatial sur un graphe discret: Pour chaque site du graphe, on tire la loi de reproduction et la loi du déplacement d'une particule vers les autres sites de manière i.i.d. dans un ensemble de lois prédéfini. Cet environnement reste ensuite fixé au cours du temps.

Superprocessus catalytique: On suppose que dans une région aléatoire, le mécanisme de branchement du superprocessus est modifié et le branchement est par exemple accéléré (ou simplement interdit dans son complémentaire). Cette région peut elle-même évoluer, si par exemple il s'agit du support d'un second superprocessus (on peut imaginer que le premier a besoin de la présence du second pour brancher, d'où l'adjectif « catalytique » en référence à des systèmes chimiques).

Les résultats associés sont des critères d'extinction (locale ou globale), le comportement asymptotique de la taille de la population ou des sous-populations à chaque site, etc. Puisque l'environnement est lui-même aléatoire, on peut poser toutes ces questions à réalisation fixée de l'environnement (approche dite *quenched*), ou en moyennant sur toutes les réalisations possibles (approche *annealed*). On consultera par exemple le chapitre VI.5 de [**AN72**] pour le premier point, [**GdH92**] et les références qui y sont données pour le second, et enfin le chapitre 9.4 de [**Eth00**] pour le troisième type de modèle.

D'une manière un peu similaire au troisième point évoqué ci-dessus, nous allons nous placer dans un espace continu, $\mathbb{R}^d$, et choisir comme région aléatoire où le comportement du processus de branchement est modifié une union de boules de la forme

$$\Gamma = \bigcup_{x \in \mathcal{P}} \overline{B}(x,r), \tag{1.22}$$

où $\overline{B}(x,r)$ est la boule fermée de centre x et de rayon $r > 0$ et $\mathcal{P}$ est un processus ponctuel de Poisson sur $\mathbb{R}^d$ d'intensité μ. Rappelons en passant la définition d'un processus ponctuel de Poisson.

DÉFINITION 1.19. *Soit μ une mesure σ-finie sur $\mathbb{R}^d$. Un processus ponctuel de Poisson $\mathcal{P}$ d'intensité μ est une variable aléatoire à valeurs dans les ensembles dénombrables de points de $\mathbb{R}^d$ telle que si $N(A)$ est le nombre de points de $\mathcal{P}$ qui appartiennent au borélien $A \subset \mathbb{R}^d$, alors*

- *pour tout A, $N(A)$ suit une loi de Poisson de paramètre $\mu(A)$;*
- *pour tous $A_1, \dots, A_k$ disjoints, $N(A_1), \dots, N(A_k)$ sont des v.a. indépendantes.*

On appelle *obstacles* les boules qui constituent Γ, car les particules qui y pénètrent sont tuées selon l'un des deux mécanismes suivants.

Obstacles durs: la loi du déplacement d'une particule est celle d'un mouvement brownien tué instantanément dès qu'il entre dans l'un des obstacles.

Obstacles mous: le déplacement d'une particule est donné par la loi d'un mouvement brownien tué à taux $a > 0$ dans Γ.

Observons qu'une particule se déplaçant parmi des obstacles mous a une probabilité positive de traverser un obstacle sans être tuée, alors que cette probabilité est nulle si elle évolue parmi des obstacles durs.

Pour motiver les questions abordées dans les chapitres suivants, commençons par considérer un seul mouvement brownien tué dans les obstacles. Notons $\mathbf{P}$ la mesure de probabilité sur l'espace probabilisé $\mathbf{\Omega}$ sur lequel Γ est défini. On rappelle que P_x est la mesure de probabilité sur l'espace de définition du mouvement brownien ξ sur $\mathbb{R}^d$ (différent de $\mathbf{\Omega}$) sous laquelle ξ part du point x.

1.3.1. *Mouvement brownien parmi des obstacles durs.* Supposons que l'on travaille avec des obstacles durs, en dimension $d \geq 2$ (la dimension 1 a relativement peu d'intérêt). Fixons $\omega \in \mathbf{\Omega}$ et notons T le temps auquel ξ entre dans $\Gamma(\omega)$, i.e.,

$$T = \inf\{t \geq 0 \ : \ \xi_t \in \Gamma(\omega)\}.$$

T dépend bien sûr de ω, même si on ne le fait pas apparaître dans la notation. Pour tout $0 \leq s \leq t$, notons $S_r(s,t)$ la *saucisse de Wiener* de rayon r le long de l'intervalle $[s,t]$, définie comme le r-voisinage de la trajectoire du mouvement brownien entre les instants s et t :

$$S_r(s,t) := \big\{y \in \mathbb{R}^d \ : \ \inf_{s \leq u \leq t} |\xi_u - y| \leq r\big\} = \bigcup_{u \in [s,t]} \big(\xi_u + \overline{B}(0,r)\big).$$

De par la structure de Γ, on a pour tout $t \geq 0$ (toujours à ω fixé)

$$\mathrm{P}_0[T > t] = \mathrm{P}_0\big[S_r(0,t) \cap \mathcal{P}(\omega) = \emptyset\big],$$

et donc par le théorème de Fubini et la première propriété énoncée dans la Définition 1.19, on obtient

$$\mathbf{E}\big[\mathrm{P}_0[T > t]\big] = \mathrm{E}_0\big[\mathbf{P}\big[S_r(0,t) \cap \mathcal{P}(\omega) = \emptyset\big]\big] = \mathrm{E}_0\big[\exp\big\{ - \mu\big(S_r(0,t)\big)\big\}\big]. \tag{1.23}$$

Dans le cas où $\mu = c\lambda$, $c > 0$ et λ étant la mesure de Lebesgue sur $\mathbb{R}^d$, on voit alors apparaître le volume de $S_r(0,t)$ dans le troisième terme de (1.23). Plus généralement, si μ admet une densité $c(x)$ par rapport à la mesure de Lebesgue, c'est $\int_{S_r(0,t)} c(x)dx$ qui dicte la probabilité *annealed* que le mouvement brownien ne soit pas encore rentré dans un obstacle au temps t.

Justement, le comportement asymptotique de $\lambda(S_\varepsilon(0,t))$ lorsque $\varepsilon \to 0$ a été l'objet d'une attention considérable, due en particulier à ses connexions avec les points multiples du mouvement brownien. Posons

$$s_d(\varepsilon) := \begin{cases} \log(\varepsilon^{-1}) & \text{si } d = 2, \\ \varepsilon^{2-d} & \text{si } d \geq 3. \end{cases}$$

Kesten, Spitzer et Whitman ont montré (cf. [**IM65**], p.253) que si $d \geq 3$,

$$\lim_{\varepsilon \to 0} s_d(\varepsilon)\lambda\big(S_\varepsilon(0,t)\big) = k_d t \qquad \mathrm{P}-\text{p.s.},$$

où $k_d = (d-2)\pi^{d/2}/\Gamma(d/2)$ ($k_3 = 2\pi$) est la capacité newtonienne de la boule unité. Le même résultat est vrai en dimension 2, avec $k_2 = \pi$ (voir [**LG86**]).

Fixons donc $c > 0$ et définissons pour tout $\varepsilon \in (0,1/2)$ la configuration d'obstacles

$$\Gamma_\varepsilon := \bigcup_{x \in \mathcal{P}^\varepsilon} \overline{B}(x,\varepsilon),$$

où $\mathcal{P}^\varepsilon$ est un processus ponctuel de Poisson d'intensité $cs_d(\varepsilon)\lambda$. Les obstacles sont ainsi de plus en plus petits, mais de plus en plus nombreux. D'après (1.23), on a alors

$$\mathbf{E}\big[\mathrm{P}_0[T > t]\big] = \mathrm{E}_0\Big[e^{-cs_d(\varepsilon)\lambda(S_\varepsilon(0,t))}\Big] \to e^{-ck_d t} \qquad \text{lorsque } \varepsilon \to 0. \tag{1.24}$$

Informellement, on peut interpréter la convergence dans (1.24) de la manière suivante : les obstacles deviennent denses dans $\mathbb{R}^d$ et le mouvement brownien devient absorbé dans les obstacles (et donc tué) à taux ck_d. Ce résultat suggère alors de s'intéresser à la convergence en loi du mouvement brownien tué instantanément dans Γ_ε vers un mouvement brownien (indépendant de tout obstacle) tué de manière homogène à taux ck_d. Un bon contrôle de la variance de $\lambda(S_\varepsilon(0,t))$ permet à Le Gall de montrer dans [**LG**] que pour tout domaine $D \subset \mathbb{R}^d$ borné et toute fonction f continue et bornée sur D,

$$\int_D \big(Q_t^{\omega,\varepsilon} f(x) - Q_t^* f(x)\big)^2 dx \to 0 \qquad \text{lorsque } \varepsilon \to 0 \tag{1.25}$$

en $\mathbf{P}$-probabilité, où $Q_t^{\omega,\varepsilon}$ (resp., Q_t^*) est le semi-groupe de transition au temps t du mouvement brownien tué en dehors de D et dans $\Gamma(\omega)$ (resp., à taux ck_d). On pourra consulter [**LG86, LG88**] pour une étude précise des variations de $\lambda(S_\varepsilon(0,t))$.

Dans le chapitre 2, on se pose la même question pour le super-mouvement brownien tué dans Γ_ε. On a alors une collection infinie de particules browniennes, qui en outre se reproduisent et meurent à cause des branchements. On obtiendra une convergence beaucoup plus forte, sous une hypothèse plus générale sur l'intensité de $\mathcal{P}^\varepsilon$. Ces résultats sont décrits dans le paragraphe 1.4.

1.3.2. *Mouvement brownien parmi des obstacles mous.* Contrairement au paragraphe précédent, si $\mu = c\lambda$ ce n'est plus le temps d'atteinte des obstacles qui dicte la probabilité de survie d'un mouvement brownien tué dans des obstacles mous, mais le temps passé par ξ dans Γ. En effet, d'après la description du mécanisme de meurtre, on a à $\omega \in \mathbf{\Omega}$ fixé

$$\mathrm{P}_0[T > t] = \mathrm{E}_0\bigg[\exp\bigg\{-a\int_0^t \mathbb{I}_\Gamma(\xi_s)ds\bigg\}\bigg],$$

où T est le temps auquel la particule brownienne est tuée. Les principaux résultats sur ce type de question sont ceux de Sznitman, dont la *méthode d'agrandissement des obstacles* (cf. [**Szn98**]) permet d'obtenir les asymptotiques suivantes :

THÉORÈME 1.20. (Théorèmes 4.5.1 et 4.5.3 de [**Szn98**])

1. *Il existe un ensemble de configurations d'obstacles de* $\mathbf{P}$*-probabilité* 1 *sur lequel*
$$\mathrm{E}_0\bigg[\exp\bigg\{-a\int_0^t \mathbb{I}_\Gamma(\xi_s)ds\bigg\}\bigg] = \exp\Big\{-C(d,c)\,\frac{t}{(\log t)^{2/d}}\,(1+o(1))\Big\}$$
lorsque $t \to \infty$*, où* $C(d,c) > 0$ *est une constante connue indépendante de* a*.*
2. *Lorsque* $t \to \infty$*,*
$$\mathbf{E}\otimes\mathrm{E}_0\bigg[\exp\bigg\{-a\int_0^t \mathbb{I}_\Gamma(\xi_s)ds\bigg\}\bigg] = \exp\big\{-C'(d,c)\,t^{\frac{d}{d+2}}\,(1+o(1))\big\},$$
pour une constante $C'(d,c) > 0$ *connue et indépendante de* a*.*

On remarquera que les asymptotiques ne sont pas les mêmes selon que l'on considère une configuration d'obstacles fixée ou la version moyennée.

Ces résultats sont du type grandes déviations : ils montrent que la probabilité qu'un mouvement brownien survive longtemps à un meurtre d'intensité constante dans les obstacles décroît exponentiellement. Dans le chapitre 3 en revanche, nous ferons tendre le taux de meurtre a vers 0 et nous serons ainsi amenés à considérer des intégrales de la forme

$$a\int_0^{a^{-1}t} \mathbb{I}_\Gamma(\xi_t)dt, \qquad a \to 0.$$

Ce sont alors leur comportement moyen et leurs fluctuations autour de leur espérance qui nous intéresserons.

1.3.3. *Résultats existants sur les processus de branchement spatiaux tués dans des obstacles.* Le Théorème 1.20 du paragraphe précédent permet à Engländer d'obtenir dans [**Eng08**] le comportement lorsque $t \to \infty$ de la masse totale d'un mouvement brownien branchant surcritique évoluant parmi des obstacles mous du type (1.22). Dans son modèle, les particules branchent en produisant exactement 2 particules partout, mais le taux de branchement est $\beta_2 > 0$ en dehors des obstacles et $\beta_1 \in (0, \beta_2)$ dans Γ. Il montre que l'espérance de la masse totale croît exponentiellement en $\beta_2 t$, avec une correction due aux obstacles en $t/(\log t)^{2/d}$ pour le résultat *quenched*, et $t^{d/(d+2)}$ pour celui *annealed*. Remarquons que la masse totale ici n'est pas un CSBP (ni même un processus markovien), car le taux de reproduction dépend de la localisation des particules et donc (1.9) appliqué à des fonctions constantes ne perd pas sa dépendance en l'espace.

Il s'intéresse aussi au taux de croissance local (dans un ouvert borné de $\mathbb{R}^d$) et à la vitesse radiale du support (i.e., la vitesse de croissance du rayon du support) et obtient $\mathbf{P}$-p.s. le même comportement asymptotique en $C(\beta_2)t$ que pour le processus sans obstacle, avec un terme correctif similaire à celui apparaissant pour la masse totale. Son approche par des arguments analytiques fournit également des résultats pour d'autres diffusions branchantes.

Dans d'autres travaux, Engländer étudie le comportement asymptotique (en temps) de mouvements browniens branchants parmi des obstacles durs dont l'intensité peut varier spatialement. Il suppose que tout le processus est tué dès que l'une des particules touche un obstacle et obtient une expression de la décroissance de la probabilité de survie en termes de solution d'un problème variationnel. Là encore, il considère un branchement surcritique et des obstacles fixés et obtient des résultats de type grandes déviations, ce qui diffère des phénomènes d'homogénéisation que l'on observera dans les chapitres 2 et 3 pour des mécanismes de branchement critiques. On renvoie à [**Eng07**] pour une revue de ces résultats.

1.4. Résultats du chapitre 2 : convergence *quenched* de super-mouvements browniens parmi des obstacles durs. Dans le chapitre 2, on se place en dimension $d \geq 2$ et on s'intéresse au comportement d'un super-mouvement brownien parmi les obstacles introduits dans le paragraphe 1.3.1. Rappelons que la mesure de probabilité $\mathbf{P}$ correspond aux obstacles, $\mathbb{P}$ au super-mouvement brownien et P à un seul mouvement brownien.

Fixons un domaine D de $\mathbb{R}^d$ et une fonction $c : \mathbb{R}^d \to \mathbb{R}_+$ mesurable et bornée. Pour tout $\varepsilon \in (0, 1/2)$, notons X^ε le super-mouvement brownien de mécanisme de branchement $\psi(u) = u^2$ et de déplacement sous-jacent la loi du mouvement brownien dans $\mathbb{R}^d$ tué à sa sortie de D et dans la collection d'obstacles durs Γ_ε définie par

$$\Gamma_\varepsilon = \bigcup_{x \in \mathcal{P}^\varepsilon} \overline{B}(x, \varepsilon),$$

où $\mathcal{P}^\varepsilon$ est un processus ponctuel de Poisson sur $\mathbb{R}^d$ d'intensité $s_d(\varepsilon)c(x)dx$. Pour rappel, $s_d(\varepsilon) = \log(\varepsilon^{-1})$ si $d = 2$ et $s_d(\varepsilon) = \varepsilon^{2-d}$ si $d \geq 3$. L'intensité des obstacles tend alors vers $+\infty$ lorsque $\varepsilon \to 0$, mais leur rayon tend vers 0. On appellera l'ensemble $\{\Gamma_\varepsilon,\ \varepsilon \in (0, 1/2)\}$ un *environnement*.

Comme le suggèrent les résultats énoncés dans le paragraphe 1.3.1, définissons également X^* comme le super-mouvement brownien de mécanisme de branchement ψ et de déplacement sous-jacent la loi du mouvement brownien tué instantanément en dehors de D et à taux $k_d c(x)$ dans D (la constante k_d est définie au paragraphe 1.3.1). Nous pouvons maintenant énoncer le résultat principal du chapitre 2.

THÉORÈME 1.21. *On a les convergences en loi suivantes :*

1. *Soit* $(\varepsilon_n)_{n\in\mathbb{N}}$ *une suite décroissante telle que*
$$\sum_n |\log\varepsilon_n|^{-1} < \infty \ \text{ si } \ d=2 \qquad \text{et} \qquad \sum_n \varepsilon_n|\log\varepsilon_n| < \infty \ \text{ si } \ d\geq 3.$$
Alors pour tout $\mu\in\mathcal{M}_f(D)$, $\mathbf{P}$*-p.s. on a*
$$\mathbb{P}_\mu(X^{\varepsilon_n})^{-1} \Rightarrow \mathbb{P}_\mu(X^*)^{-1},$$
où on a noté $\mathbb{P}Z^{-1}$ *la loi de la variable* Z *sous la mesure de probabilité* $\mathbb{P}$.

2. *Pour tout* $\mu\in\mathcal{M}_f(D)$, *lorsque* $\varepsilon\to 0$ *la loi de* X^ε *sous* $\mathbb{P}_\mu$ *converge en* $\mathbf{P}$*-probabilité vers la loi de* X^* *sous* $\mathbb{P}_\mu$.

Le premier résultat est donc une convergence en loi *quenched* le long d'une sous-suite (ε_n) tendant suffisamment rapidement vers 0 ; le second nous dit que la distance (de Prohorov) entre les lois de X^ε et de X^* tend vers 0 en $\mathbf{P}$-probabilité, cette fois sans condition de décroissance sur ε.

Comme on l'a vu au paragraphe 1.2.3, la preuve du point 1. du Théorème 1.21 procède en deux étapes. Pour montrer la convergence des marginales de dimension finie de X^{ε_n} énoncée dans la Proposition 2.6 du chapitre 2, on utilise leur caractérisation donnée dans la Remarque 1.8 du paragraphe 1.2.1. Fixons donc $p\in\mathbb{N}$ et $0\leq t_1<\ldots<t_p$. Pour toutes fonctions $f_1,\ldots,f_p\in\mathcal{B}_b^+(\mathbb{R}^d)$, on a
$$\mathbb{E}_\mu\Big[\exp\Big\{-\sum_{i=1}^p\langle X^\varepsilon_{t_i},f_i\rangle\Big\}\Big] = e^{-\langle\mu,w_0^\varepsilon\rangle},$$
$$\mathbb{E}_\mu\Big[\exp\Big\{-\sum_{i=1}^p\langle X^*_{t_i},f_i\rangle\Big\}\Big] = e^{-\langle\mu,w_0^*\rangle},$$
où si on note T le temps de sortie de D du mouvement brownien ξ et T_ε son temps d'entrée dans Γ_ε, les fonctions w^ε et w^* sont les uniques solutions positives des équations intégrales suivantes :
$$w_t^\varepsilon(x) + \mathrm{E}_{t,x}\Big[\int_t^\infty w_s^\varepsilon(\xi_s)^2\,\mathbb{I}_{\{s<T\wedge T_\varepsilon\}}ds\Big] = \sum_{i=1}^p \mathrm{E}_{t,x}\Big[f_i(\xi_{t_i})\,\mathbb{I}_{\{t_i<T\wedge T_\varepsilon\}}\Big],$$
$$w_t^*(x) + \mathrm{E}_{t,x}\Big[\int_t^\infty w_s^*(\xi_s)^2 e^{-k_d\int_t^s c(\xi_u)du}\,\mathbb{I}_{\{s<T\}}ds\Big] = \sum_{i=1}^p \mathrm{E}_{t,x}\Big[f_i(\xi_{t_i})e^{-k_d\int_t^{t_i} c(\xi_u)du}\,\mathbb{I}_{\{t_i<T\}}\Big].$$
En intégrant sur D la différence entre ces deux équations, on obtient une inégalité de la forme (1.15), où le supremum est remplacé par une intégrale sur D (voir l'équation (2.19)). La difficulté majeure réside alors dans la preuve que le terme de reste converge bien $\mathbf{P}$-p.s. vers 0 si on se restreint à la sous-suite $(\varepsilon_n)_{n\in\mathbb{N}}$, ce qui découle d'estimées sur des intégrales de fonctions mesurables et bornées sur la saucisse de Wiener établies dans le paragraphe 2 du chapitre 2. Un argument de type Gronwall permet ensuite d'en déduire que $\mathbf{P}$-p.s., $\int_D|w_t^{\varepsilon_n}(x)-w_t^*(x)|dx$ converge vers 0 sur un ensemble de temps inclus dans $[0,t_p]$ dont le complémentaire dans $[0,t_p]$ est de mesure de Lebesgue nulle, puis plusieurs applications du théorème de Fubini nous donnent la convergence
$$e^{-\langle\mu,w_0^{\varepsilon_n}(\cdot)\rangle}\to e^{-\langle\mu,w_0^*(\cdot)\rangle}, \qquad \mathbf{P}-\text{p.s.}$$
On conclut en considérant un ensemble dénombrable de p-uplets (p variant dans $\mathbb{N}$) de fonctions $(f_1,\ldots,f_p)$ tel que la famille correspondante d'applications $(\mu_1,\ldots,\mu_p)\mapsto\exp\left\{-\sum_{i=1}^p\langle\mu_i,f_i\rangle\right\}$ détermine la convergence.

Ensuite, on définit un *bon environnement* comme étant un environnement pour lequel toutes les marginales fini-dimensionnelles de X^{ε_n} indexées par des temps rationnels convergent vers celles de X^*. L'ensemble des bons environnements est donc de $\mathbf{P}$-probabilité 1 d'après le résultat précédent. On montre alors dans le paragraphe 4 du chapitre 2 que dans un bon environnement la suite $(X^{\varepsilon_n})_{n\in\mathbb{N}}$ est tendue, en utilisant les résultats énoncés dans le paragraphe 1.2.3 et différentes bornes sur les variations des processus à valeurs réelles de la forme $\langle X^{\varepsilon_n}, \varphi\rangle$. Le point 2. du Théorème 1.21 est un corollaire immédiat du point 1.

1.5. Résultats du chapitre 3 : mouvements browniens branchants parmi des obstacles mous. On considère à nouveau des obstacles aléatoires dans $\mathbb{R}^d$ ($d \geq 1$), définis de manière un peu plus générale que dans les paragraphes précédents. Pour un $r_0 > 0$ fixé, on note $\mathcal{K}_{r_0}$ l'ensemble des compacts de $\mathbb{R}^d$ inclus dans la boule fermée $\overline{B}(0, r_0)$ (muni de la distance de Hausdorff). Fixons une mesure finie Θ sur $\mathcal{K}_{r_0}$ et définissons une mesure ponctuelle de Poisson

$$\mathcal{N} := \sum_{i\in I} \delta_{(x_i,K_i)},$$

sur $\mathbb{R}^d \times \mathcal{K}_{r_0}$, d'intensité $\lambda \otimes \Theta$. La collection d'obstacles Γ_ω est alors donnée pour tout $\omega \in \mathbf{\Omega}$ par

$$\Gamma_\omega = \bigcup_{i\in I} (x_i + K_i).$$

Notons κ la constante

$$\kappa := \mathbf{P}[0 \in \Gamma_\omega] = 1 - \exp\Big(- \int_{\mathcal{K}_{r_0}} \lambda(K)\Theta(dK)\Big),$$

que nous supposerons strictement positive dans la suite. La loi des obstacles étant invariante par translation, on a également $\mathbf{P}[x \in \Gamma_\omega] = \kappa$ pour tout $x \in \mathbb{R}^d$.

Pour tout $\omega \in \mathbf{\Omega}$ et $\varepsilon \in (0, 1)$, définissons un mouvement branchant $Z^{\omega,\varepsilon}$ de la manière suivante : chaque particule, indépendamment des autres,

- se déplace dans $\mathbb{R}^d$ suivant la loi d'un mouvement brownien tué à taux ε dans Γ_ω ;
- branche à taux 1 et produit un nombre de descendants de loi fixée et commune à toutes les particules, d'espérance 1 et de variance $\sigma^2 > 0$ finie.

On s'intéresse alors à l'influence à long terme des obstacles sur le comportement du mouvement brownien branchant. Plus précisément, on cherche à répondre à la question suivante : si on fixe un domaine $A \subset \mathbb{R}^d$ contenant l'origine et que l'on fait partir notre mouvement brownien branchant d'une seule particule en 0, quelle est la probabilité que l'une des particules de $Z^{\omega,\varepsilon}$ atteigne le complémentaire du domaine RA, pour une constante $R \gg 1$? En l'absence d'obstacles, des résultats classiques assurent que cette probabilité décroît en CR^{-2} lorsque $R \to \infty$, pour une constante $C > 0$. Lorsque les particules du mouvement brownien branchant sont tuées à taux faible dans les obstacles, la réponse dépend de l'ordre de grandeur de εR^2. En effet, le résultat principal du chapitre 3 est le théorème suivant :

THÉORÈME 1.22. *Pour tout $a \geq 0$, notons $u_{(a)} = (u_{(a)}(x), x \in A)$ l'unique solution positive du système*

$$\begin{cases} \frac{1}{2}\Delta u = \frac{\sigma^2}{2}u^2 + au & \text{dans } A, \\ u|_{\partial A} = +\infty. \end{cases}$$

Alors

$$\lim_{R\to\infty} \Big(\sup_{\varepsilon\geq 0} \Big| R^2 \mathbb{P}_{\delta_0}(Z^{\omega,\varepsilon} \text{ atteint } (RA)^c) - u_{(\kappa\varepsilon R^2)}(0) \Big| \Big) = 0, \qquad \mathbf{P}(d\omega)\text{–a.s.}$$

Ainsi, sauf pour un ensemble de $\mathbf{P}$-probabilité nulle de réalisations des obstacles, lorsque $\varepsilon R^2 \to a \in [0, \infty)$ la probabilité que l'une des particules de $Z^{\omega,\varepsilon}$ atteigne le complémentaire de RA décroît en $u_{(\kappa a)}(0)/R^2$ (la continuité de la fonction $b \mapsto u_{(b)}(0)$ est établie dans le Lemme 3.7 du chapitre 3). Notons que si $a < \infty$, on a bien $u_{(\kappa a)} > 0$ et donc la probabilité d'atteinte de

$(RA)^c$ décroît à une vitesse comparable à celle que l'on obtient en l'absence de meurtre dans les obstacles. L'influence des obstacles n'apparaît que dans la constante multiplicative en facteur de R^{-2}, et uniquement à travers le produit εR^2 et le paramètre κ défini précédemment.

En revanche, $u_{(a)}(0) \to 0$ lorsque $a \to \infty$ et donc le Théorème 1.22 ne donne que peu d'informations sur la manière dont la probabilité d'atteinte de $(RA)^c$ décroît lorsque $\varepsilon R^2 \to \infty$. Dans la Proposition 3.2, on montre en fait qu'il existe deux constantes $\alpha, \beta > 0$ ne dépendant que de la mesure Θ et du mécanisme de branchement des particules, telles que pour tous $R \geq 1$ et $\varepsilon \in (0,1)$,

$$\varepsilon \ \exp(-\alpha R\sqrt{\varepsilon}) \leq \mathbf{P} \otimes \mathbb{P}_{\delta_0}(Z^{\omega,\varepsilon} \text{ atteint } (RA)^c) \leq \exp(-\beta R\sqrt{\varepsilon}).$$

La preuve du Théorème 1.22 repose sur un résultat d'homogénéisation *quenched* que nous décrivons à présent. A ω et ε fixés, définissons pour tout $t \geq 0$ une nouvelle mesure aléatoire $X_t^{\omega,\varepsilon}$ par : pour toute fonction φ mesurable bornée,

$$\langle X_t^{\omega,\varepsilon}, \varphi\rangle = \varepsilon \int_{\mathbb{R}^d} Z^{\omega,\varepsilon}_{\varepsilon^{-1}t}(dx)\varphi\big(\sqrt{\varepsilon}x\big).$$

Par la propriété d'échelle du mouvement brownien, il n'est pas difficile de voir que $\varepsilon^{-1}X^{\omega,\varepsilon}$ est un mouvement brownien branchant dont les particules branchent à taux ε^{-1} et se déplacent selon la loi d'un mouvement brownien tué à taux 1 dans $\sqrt{\varepsilon}\,\Gamma_\omega$. Lorsque $\varepsilon \to 0$, les obstacles renormalisés deviennent denses dans $\mathbb{R}^d$ et les particules deviennent tuées de manière homogène à taux κ dans $\mathbb{R}^d$. On a alors le résultat suivant, correspondant au Théorème 3.3 du chapitre 3. On note $\lfloor z\rfloor$ la partie entière de $z \in \mathbb{R}$.

THÉORÈME 1.23. *Presque sûrement pour la loi des obstacles, la loi de* $(X_t^{\omega,\varepsilon})_{t\geq 0}$ *sous* $\mathbb{P}_{\lfloor\varepsilon^{-1}\rfloor\delta_0}$ *converge lorsque* $\varepsilon \to 0$ *vers celle du super-mouvement brownien* $Y^{(\kappa)}$ *de mécanisme de branchement* $\psi_{(\kappa)}(u) = \frac{\sigma^2}{2}u^2 + \kappa u$ *et de valeur initiale* δ_0.

On montre ensuite dans le Lemme 3.13 que si $R\sqrt{\varepsilon} \to b$, la probabilité d'atteinte de $(R\sqrt{\varepsilon}A)^c$ par $X^{\omega,\varepsilon}$ converge $\mathbf{P}$-presque sûrement vers la probabilité d'atteinte de $(bA)^c$ par $Y^{(\kappa)}$, ce qui nous permet de conclure la preuve du Théorème 1.22.

2. Généalogies de populations structurées spatialement

Passons maintenant à la seconde partie de notre étude, à savoir les *processus de coalescence spatiaux* et leurs applications à la modélisation de généalogies de populations. Dans la suite, pour tout $k \in \mathbb{N}$ on notera $[k]$ l'ensemble $\{1,\dots,k\}$, $\mathcal{P}_k$ l'ensemble des partitions de $[k]$ et $\mathcal{P}_\infty$ l'ensemble des partitions de $\mathbb{N}$. Enfin, si $\pi \in \bigcup_n \mathcal{P}_n$, on notera $|\pi|$ le nombre de blocs de π.

2.1. Coalescents échangeables et liens avec certains modèles de population. Commençons par définir les *coalescents échangeables*, que nous retrouverons comme objets limites des processus spatiaux étudiés dans les chapitres 4 et 5. Il s'agit de processus markoviens à valeurs dans $\mathcal{P}_\infty$ dont le mécanisme d'évolution n'autorise que des fusions de blocs, d'où l'appellation « coalescents » ; « échangeables » signifie que leur évolution ne dépend pas des étiquettes contenues dans chaque bloc, que l'on peut donc permuter (échanger) au départ sans modifier la loi du processus. Nous avons opté ici pour une présentation par ordre croissant de complexité, mais soulignons que le coalescent de Kingman et les Λ-coalescents définis ci-dessous sont des cas particuliers de Ξ-coalescents.

2.1.1. *Définitions.* Pour tous $n, p \in \mathbb{N}$, $k_1, \dots, k_p \geq 2$ et $\zeta, \eta \in \mathcal{P}_n$, notons $\eta \subset_{k_1,\dots,k_p} \zeta$ si η est obtenue à partir de ζ en fusionnant exactement k_1 blocs de ζ en un seul (plus gros) bloc,

k_2 autres blocs de ζ en un bloc différent de η, ... , et en gardant les autres blocs inchangés. Par exemple,

$$\big\{\{1\},\{2,3,4\},\{5\},\{6,7\}\big\} \subset_{2,3} \big\{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\},\{7\}\big\}.$$

Appelons *restriction à $\mathcal{P}_k$ d'une partition $\pi \in \mathcal{P}_{k'}$, $k' \in \{k+1,\ldots,\infty\}$* l'élément de $\mathcal{P}_k$ obtenu en « effaçant » les étiquettes $k+1,\ k+2,\ldots$, de π. On note $\pi\big|_{\mathcal{P}_k}$ cette nouvelle partition. Par exemple,

$$\big\{\{1\},\{2,4,7\},\{3,8\},\{5\},\{6,9\}\big\}\Big|_{\mathcal{P}_5} = \big\{\{1\},\{2,4\},\{3\},\{5\}\big\}.$$

Par extension, on appelle *restriction à $\mathcal{P}_k$ d'un processus* $\{\Pi_t, t\geq 0\}$ à valeurs dans $\mathcal{P}_{k'}$ le processus Π' (ou $\Pi\big|_{\mathcal{P}_k}$) défini par

$$\Pi'_t := \Pi_t\big|_{\mathcal{P}_k}, \qquad \forall\, t \geq 0.$$

Le premier coalescent que nous allons définir a été introduit par Kingman dans [**Kin82**].

DÉFINITION 1.24. (Coalescent de Kingman) *Le coalescent de Kingman est le processus markovien à valeurs dans $\mathcal{P}_\infty$ tel que pour tout $n\in\mathbb{N}$, sa restriction à $\mathcal{P}_n$ est un processus markovien dont les taux de transition q_K sont donnés par : si $\zeta \neq \eta \in \mathcal{P}_n$,*

$$q_K(\zeta,\eta) = \begin{cases} 1 & \text{si } \eta \subset_2 \zeta, \\ 0 & \text{sinon.} \end{cases}$$

En d'autres termes, on autorise la fusion de seulement deux blocs à la fois et chacune des $\binom{|\zeta|}{2}$ paires de blocs essaie de fusionner à taux 1 jusqu'à ce qu'une première paire y arrive. Le temps à attendre pour observer cette première fusion est donc une variable aléatoire exponentielle de paramètre $\binom{|\zeta|}{2}$ et puisque toutes les paires fusionnent au même taux la nouvelle valeur du coalescent est tirée uniformément au hasard dans l'ensemble des partitions η telles que $\eta \subset_2 \zeta$.

Le type de coalescent suivant a été introduit indépendamment par Pitman dans [**Pit99**] et par Sagitov dans [**Sag99**]. On donne ici la caractérisation utilisée par Pitman, dont on présentera une interprétation dans la Remarque 1.28.

DÉFINITION 1.25. (Λ-coalescent) *Soit Λ une mesure finie sur* $[0,1]$. *On appelle* coalescent à collisions multiples de mesure Λ, *ou* Λ-coalescent, *le processus markovien à valeurs dans $\mathcal{P}_\infty$ tel que pour tout $n\in\mathbb{N}$, sa restriction à $\mathcal{P}_n$ est un processus markovien dont les taux de transition q_Λ sont donnés par : si $\zeta,\eta\in\mathcal{P}_n$, $b := |\zeta|$ et $\eta \subset_k \zeta$ pour un $k\in\{2,\ldots,b\}$, alors*

$$q_\Lambda(\zeta,\eta) = \int_0^1 x^k(1-x)^{b-k}\,\frac{\Lambda(dx)}{x^2}, \tag{1.26}$$

et les autres taux de transition (pour $\zeta \neq \eta$) sont nuls.

Observons que si $b=2$, nécessairement $k=2$ et les deux blocs fusionnent à taux $\Lambda([0,1]) < \infty$. Il n'est pas difficile de voir que ce taux est un majorant de tous les autres, qui sont donc également finis. Par ailleurs, on retrouve le coalescent de Kingman en prenant $\Lambda = \delta_0$, la masse de Dirac en 0. Sauf dans ce cas particulier, on autorise des fusions de plus de 2 blocs en un seul plus gros bloc, mais pas de coalescence simultanée en plusieurs blocs distincts. C'est ce dernier point qui conduit à la définition des coalescents échangeables les plus généraux (parmi ceux possédant la Propriété 1.27 ci-dessous de *cohérence*), introduits par Möhle et Sagitov dans [**MS01**] et par Schweinsberg dans [**Sch00**]. Notons Δ le simplexe

$$\Delta := \Big\{\mathbf{x} = (x_1,x_2,\ldots)\ :\ x_1 \geq x_2 \geq \ldots \geq 0,\ \sum_{i=1}^{\infty} x_i \leq 1\Big\}.$$

DÉFINITION 1.26. (Ξ-coalescent) *Soit* Ξ *une mesure finie sur* Δ*, que l'on peut (toujours) écrire* $\Xi = \Xi_0 + a\delta_{\mathbf{0}}$ *avec* $a \geq 0$ *et* Ξ_0 *ne possèdant pas d'atome en zéro. On appelle* coalescent à collisions multiples et simultanées de mesure Ξ*, ou* Ξ-coalescent*, le processus markovien à valeurs dans* $\mathcal{P}_\infty$ *tel que pour tout* $n \in \mathbb{N}$*, sa restriction à* $\mathcal{P}_n$ *est un processus markovien dont les taux de transition* q_Ξ *sont donnés par : si* $\zeta, \eta \in \mathcal{P}_n$*,* $\eta \subset_{k_1,\ldots,k_r} \zeta$ *avec* $k_1, \ldots, k_r \geq 2$ *et si* $s := |\zeta| - \sum_{i=1}^{r} k_i$*, alors*

$$q_\Xi(\zeta,\eta) = \int_\Delta \left(\sum_{l=0}^{s} \sum_{i_1 \neq \ldots \neq i_{r+l}} \binom{s}{l} x_{i_1}^{k_1} \ldots x_{i_r}^{k_r} x_{i_{r+1}} \ldots x_{i_{r+l}} \Big(1 - \sum_{j=1}^{\infty} x_j\Big)^{s-l} \right) \frac{\Xi_0(d\mathbf{x})}{\sum_{j=1}^{\infty} x_j^2}$$
$$+a\, \mathbb{I}_{\{r=1,k_1=2\}}, \tag{1.27}$$

et les autres taux de transition (pour $\zeta \neq \eta$*) sont nuls. Puisque ces taux ne dépendent que de* $b := |\zeta|$*,* $k_1, \ldots, k_r$ *et* s*, on les note* $\lambda_{b;k_1,\ldots,k_r;s}$*.*

A nouveau, observons que si ζ contient deux blocs, le taux auquel ils fusionnent est $\lambda_{2;2;0} = \Xi_0(\Delta) + a = \Xi(\Delta) < \infty$ et on peut vérifier que ce taux est un majorant de tous les autres. Les Λ-coalescents sont des cas particuliers de Ξ-coalescents pour lesquels $\Xi(\{x_2 > 0\}) = 0$. La Figure 1 donne deux représentations d'une réalisation d'un Ξ-coalescent. On verra que celle de gauche, sous forme d'arbre, correspond très bien à l'interprétation en termes de processus généalogique d'un échantillon d'individus.

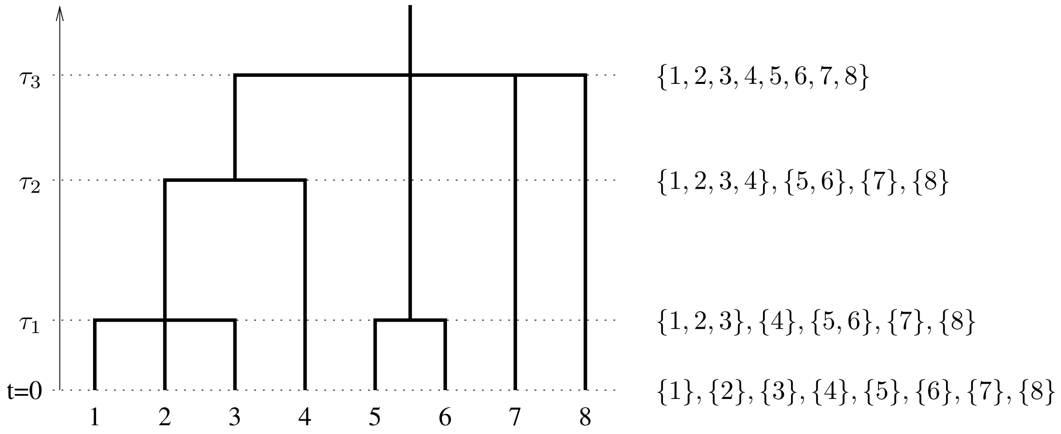

FIG. 1. Réalisation possible d'un Ξ-coalescent sur $\mathcal{P}_8$: on part de la partition triviale au temps 0, puis au temps τ_1 on observe une première transition au cours de laquelle $\{1\}, \{2\}$ et $\{3\}$ fusionnent en un seul bloc et $\{5\}, \{6\}$ en un autre. Les singletons restants sont inchangés. L'évolution continue jusqu'au temps τ_3 auquel le processus atteint la valeur finale $\{1, \ldots, 8\}$.

L'expression (1.27) semble compliquée au premier abord, mais elle se comprend très facilement grâce à la construction des coalescents à partir d'un processus ponctuel de Poisson que nous donnerons dans la Remarque 1.28. Avant cela, disons un mot de la construction de tous ces processus. Restreint à $\mathcal{P}_n$ ($n \in \mathbb{N}$), un coalescent est un processus de sauts à valeurs dans un espace fini, dont l'évolution ne dépend ni des étiquettes contenues dans les blocs, ni même de la taille de ces derniers. Seul le nombre de blocs gouverne les transitions, ce qui confère aux coalescents la propriété de *cohérence* (ou *consistance* d'après la terminologie anglaise) suivante :

PROPRIÉTÉ 1.27. (Cohérence) *Soit* Π *un* Ξ*-coalescent et* $m < n$ *deux entiers. Alors*

$$\left\{\Pi|_{\mathcal{P}_n}\right\}\Big|_{\mathcal{P}_m} \overset{(d)}{=} \Pi|_{\mathcal{P}_m}.$$

En oubliant les étiquettes $m+1,\ \dots,\ n$ dans $\Pi\big|_{\mathcal{P}_n}$, on obtient donc un processus dont la loi est celle de $\Pi\big|_{\mathcal{P}_m}$. La preuve consiste simplement à vérifier que les taux de transition des deux processus de sauts coïncident, ce qui est suffisant puisque leur espace d'états est fini. Grâce à la propriété de cohérence et au théorème d'extension de Kolmogorov, on peut donc construire un processus à valeurs dans $\mathcal{P}_\infty$ tel que nous l'avons supposé dans les définitions ci-dessus. Par ailleurs, la Proposition 25 de [**Sch00**] garantit que les Ξ-coalescents sont des processus fellerien et donc satisfont la propriété de Markov forte. On renvoie au paragraphe 5 de [**Sch00**] pour d'autres propriétés de régularité des Ξ-coalescents. De manière canonique, on choisit la partition de $\mathbb{N}$ en singletons comme valeur initiale d'un coalescent et son évolution s'arrête lorsqu'il atteint la valeur absorbante $\{\{1,2,\dots\}\}$ (s'il l'atteint).

REMARQUE 1.28. (Construction poissonienne) *Pour comprendre l'expression des taux de transition donnée dans (1.27), supposons que* $a = 0$ *(on dit que le coalescent* n'a pas de partie Kingman*) et considérons la construction suivante du* Ξ*-coalescent, directement sur* $\mathcal{P}_\infty$*, donnée dans le paragraphe 3 de* [**Sch00**]. *Soit* Σ *un processus ponctuel de Poisson sur* $\mathbb{R}_+ \times \Delta$ *d'intensité* $dt \otimes \big(\sum_i x_i^2\big)^{-1}\Xi(d\mathbf{x})$ *(on peut vérifier qu'il s'agit bien d'une mesure* σ*-finie). Si* $(t, \mathbf{x}) \in \Sigma$*, alors au temps* t *chaque bloc de la partition* Π_{t-} *indépendamment des autres choisit la couleur* i *avec probabilité* x_i *(pour tout* $i \in \mathbb{N}$*), ou reste sans couleur avec probabilité* $1 - \sum_{i=1}^\infty x_i$*. Pour former* Π_t*, on fusionne tous les blocs ayant adopté la même couleur et les autres restent inchangés.*

Pour obtenir le taux de transition $q_\Xi(\zeta, \eta)$ *donné dans (1.27), il faut donc trouver les coloriages des blocs de* ζ *qui conduisent à* η*. Observons tout d'abord que les blocs restant inchangés durant cette transition peuvent soit ne pas avoir été coloriés, soit être les seuls à avoir choisi une certaine couleur. On commence alors par déterminer l'élément* $\mathbf{x}$ *de* Δ *donnant les probabilités de chaque couleur en fonction de l'intensité* $\big(\sum_i x_i^2\big)^{-1}\Xi(d\mathbf{x})$*, puis le nombre* $l \in \{0,\dots,s\}$ *de blocs inchangés mais tout de même coloriés. Il nous faut donc* $r + l$ *couleurs différentes, d'où la somme sur* $\{i_1 \neq \dots \neq i_{r+l}\}$ *dans laquelle le terme binomial correspond au nombre de choix possibles des* l *blocs coloriés parmi les* s *qui restent inchangés. Pour finir, les* k_1 *blocs qui fusionnent en un bloc de* η *ont une probabilité* $x_{i_1}^{k_1}$ *de tous choisir la couleur* i_1*,* $x_{i_2}^{k_2}$ *est la probabilité que les* k_2 *blocs suivants qui fusionnent adoptent tous la couleur* i_2*, ... , et les derniers termes de la forme* x_{r+j}^1 *correspondent au coloriage d'un bloc inchangé par la couleur* i_{r+j}*. Les blocs inchangés restants ont quant à eux une probabilité* $\big(1 - \sum_i x_i\big)^{s-l}$ *qu'aucun d'entre eux ne soit colorié.*

Dans le cas particulier des Λ*-coalescents, les blocs ne peuvent être coloriés qu'en une seule couleur et* Λ *est en fait la distribution de* x_1*, ce qui permet de retrouver l'expression (1.26).*

Les propriétés des coalescents ont été et continuent à être largement étudiées, car ils constituent à la fois des objets mathématiques intéressants par leur richesse et leur complexité, mais également des outils à présent communément utilisés pour l'étude de la diversité génétique au sein d'une population. On renvoie à [**Be09**] pour une revue du premier point et à [**Wak08**] pour les applications en génétique des populations.

2.1.2. *Généalogies de populations non spatiales.* Avant de considérer des populations réparties sur un espace comme nous le ferons dans les chapitres 4 et 5, montrons comment on peut obtenir des coalescents échangeables comme limites de processus généalogiques de populations non-structurées (aussi appelées *panmictiques*). Les résultats présentés ici sont dus à Möhle et Sagitov [**MS01**], même si on en donne la version de Schweinsberg [**Sch00**] afin d'utiliser sa caractérisation des Ξ-coalescents (cf. Définition 1.26). La Proposition 1.31 nous servira en particulier à mettre en avant les facteurs qui limitent les processus généalogiques à des fusions binaires (i.e., conduisent à un coalescent de Kingman) ou au contraire génèrent des collisions multiples, voire simultanées.

Commençons par deux définitions.

DÉFINITION 1.29. *On dit qu'un vecteur aléatoire* $(X_1,\dots,X_n)$ *est* échangeable *si sa loi est invariante par permutation des coordonnées, i.e., si pour toute permutation* σ *de* $[n]$*,*

$$(X_{\sigma(1)},\dots,X_{\sigma(n)}) \overset{(\mathrm{d})}{=} (X_1,\dots,X_n).$$

Le modèle suivant a été introduit par Cannings dans [**Can74**].

DÉFINITION 1.30. *Fixons* $N\in\mathbb{N}$ *et considérons une population panmictique de taille constante égale à* N *et se reproduisant par générations. Soit* $\nu := (\nu_1,\dots,\nu_N)$ *un vecteur aléatoire échangeable tel que* $\sum_{k=1}^N \nu_k = N$ *p.s. et* $(\nu^i)_{i\geq 1}$ *une suite i.i.d. de vecteurs de même loi que* ν*. Le* modèle de Cannings *associé à* ν *est défini de la manière suivante : pour tout* $i\in\mathbb{N}$*, la génération* $i+1$ *est déterminée par le vecteur* ν^i*, au sens où* ν_k^i *donne le nombre de descendants de l'individu* k *de la* i*-ème génération. Les individus de la nouvelle génération sont étiquetés de manière échangeable.*

Puisque ν est échangeable, tous les individus d'une même génération ont une chance identique de se reproduire. Observons par ailleurs que la taille de la population est bien constante puisqu'on impose $\sum_{k=1}^N \nu_k = N$ p.s. Cette hypothèse et l'échangeabilité de ν assurent alors que pour tout $k\in\{1,\dots,N\}$, on a $\mathbb{E}[\nu_k]=1$.

Par commodité, supposons que la population décrite par un modèle de Cannings est également définie pour tous les temps négatifs, afin de pouvoir remonter dans le « passé » aussi loin qu'il le faut. Prenons au hasard deux individus distincts à l'instant 0 que nous appelons le présent. Si on note c_N la probabilité que ces deux individus aient le même parent à la génération précédente, on obtient facilement

$$c_N = \sum_{k=1}^N \mathbb{P}[\text{ parent commun} = k\ ] = \sum_{k=1}^N \mathbb{E}\left[\frac{\nu_k}{N}\frac{(\nu_k-1)}{(N-1)}\right] = \frac{\mathbb{E}[\nu_1(\nu_1-1)]}{N-1},$$

où la dernière égalité découle de l'échangeabilité de ν. En effet, pour que le parent commun de ces deux individus soit l'individu k de la génération précédente, il faut avoir choisi le premier parmi les ν_k descendants de k et que dans les $N-1$ individus de la génération 0 restants, le second ait été choisi parmi les ν_k-1 autres descendants de k. S'ils n'ont pas le même parent à la génération précédente, ils ont alors la probabilité c_N d'avoir le même grand-parent, et ainsi de suite jusqu'à ce que l'on arrive au premier ancêtre commun aux deux individus. Le nombre de générations à remonter dans le passé pour atteindre cet ancêtre commun est donc une variable aléatoire géométrique de probabilité de succès c_N, d'espérance c_N^{-1}.

En utilisant les mêmes arguments que pour c_N, la probabilité que n_1 individus aient le même parent à la génération précédente, n_2 autres aient un même parent distinct du premier, ... , et n_r individus aient encore un autre parent en commun (où $r\geq 1$, $n_i\geq 1$ et bien sûr $\sum_i n_i\leq N$) est donnée par

$$\begin{aligned} p_N(n_1,\dots,n_r) &= \sum_{i_1\neq\ldots\neq i_r} \mathbb{E}\left[\frac{(\nu_{i_1})_{n_1}}{(N)_{n_1}}\frac{(\nu_{i_2})_{n_2}}{(N-n_1)_{n_2}}\cdots\frac{(\nu_{i_r})_{n_r}}{(N-n_1-\ldots-n_{r-1})_{n_r}}\right] \\ &= \frac{(N)_r}{(N)_{n_1+\ldots+n_r}}\,\mathbb{E}\big[(\nu_1)_{n_1}\cdots(\nu_r)_{n_r}\big], \end{aligned}$$

où $(M)_n := M(M-1)\cdots(M-n+1)$.

Avant d'énoncer le résultat principal de ce paragraphe, formalisons la notion de processus généalogique de k individus que l'on pourra appeler 1, 2, …, k sans perte de généralité puisque les individus d'un modèle de Cannings sont étiquetés de manière échangeable. Prenons comme espace de travail $\mathcal{P}_k$ et définissons la chaîne de Markov $(R_t^{[k]})_{t\geq 0}$ à valeurs dans $\mathcal{P}_k$ par : pour tout $t \geq 0$, i et j sont dans le même bloc de $R_t^{[k]}$ ssi les individus étiquetés i et j au temps 0 ont le même ancêtre à la génération $-t$ (dans le passé). On a donc

$$R_0^{[k]} = \big\{\{1\},\{2\},\dots,\{k\}\big\},$$

et chaque bloc B de $R_t^{[k]}$ correspond à un individu vivant à la génération $-t$ dont la descendance au temps 0 contient tous les individus dont l'étiquette est dans B. Puisque chaque individu n'a qu'un seul parent à la génération précédente, $R_t^{[k]}$ est bien une partition de $[k]$ pour tout t et les blocs de $R^{[k]}$ ne peuvent que fusionner (en atteignant un ancêtre commun). On obtient donc bien un arbre généalogique de la forme de celui proposé dans la Figure 1.

Un calcul rapide montre ensuite que $R^{[k]}$ possède la Propriété 1.27 de cohérence. En effet, appelons R' la restriction de $R^{[k]}$ à $\mathcal{P}_{k-1}$ et calculons la probabilité d'une transition dans laquelle n_1 blocs de R'_t fusionnent en un bloc de R'_{t+1}, n_2 en un autre bloc, ... , et n_r en un bloc distinct des précédents (où $n_i \geq 1$ pour tout i et $\sum_{i=1}^r n_i = |R'_t|$). A partir du moment où $\{k\}$ fusionne avec un autre bloc, les probabilités de transition sont les mêmes pour R' et $R^{[k-1]}$ puisqu'elles ne dépendent pas de la taille des blocs (juste de la probabilité que les individus de la génération $-t$ correspondant à chaque bloc aient la distribution de parents appropriée à la génération $-(t+1)$). Avant cet instant, exactement $r+1$ transitions de $R^{[k]}$ génèrent la transition de R' recherchée : celles au cours desquelles $\{k\}$ fusionne également avec l'un des r groupes de blocs et celle pendant laquelle $\{k\}$ reste inchangé. Ceci nous donne comme probabilité de transition, en notant $b := |R'_t| = |R_t^{[k]}| - 1$,

$$\begin{aligned}
&\sum_{i=1}^r p_N\left(n_1,\dots,n_i+1,\dots,n_r\right) + p_N(n_1,\dots,n_r,1)\\
&\quad= \frac{(N)_r}{(N)_{b+1}}\sum_{i=1}^r \mathbb{E}\big[(\nu_1)_{n_1}\cdots(\nu_i)_{n_i+1}\cdots(\nu_r)_{n_r}\big] + \frac{(N)_{r+1}}{(N)_{b+1}}\,\mathbb{E}\big[(\nu_1)_{n_1}\cdots(\nu_r)_{n_r}\nu_{r+1}\big]\\
&\quad= \frac{(N)_r}{(N)_{b+1}}\,\mathbb{E}\Bigg[(\nu_1)_{n_1}\cdots(\nu_r)_{n_r}\bigg\{\sum_{i=1}^r(\nu_i-n_i)+(N-r)\nu_{r+1}\bigg\}\Bigg]\\
&\quad= \frac{(N)_r}{(N)_{b+1}}\,\mathbb{E}\Bigg[(\nu_1)_{n_1}\cdots(\nu_r)_{n_r}\bigg\{\sum_{i=1}^N\nu_i-\sum_{i=1}^r n_i\bigg\}\Bigg]\\
&\quad= \frac{(N)_r}{(N)_b}\,\mathbb{E}\big[(\nu_1)_{n_1}\cdots(\nu_r)_{n_r}\big] = p_N(n_1,\dots,n_r),
\end{aligned}$$

où l'avant-dernière ligne utilise l'échangeabilité de ν pour remplacer $(N-r)\nu_{r+1}$ par $\sum_{i=r+1}^N \nu_i$ et la dernière ligne découle de l'égalité $\sum_{i=1}^N \nu_i - \sum_{i=1}^r n_i = N-b$. On retrouve donc bien la probabilité de transition correspondant à $R^{[k-1]}$ et la famille de chaînes de Markov $\{R^{[k]},\ k\in\mathbb{N}\}$ satisfait la propriété de cohérence à N fixé.

Faisons maintenant tendre la taille de la population vers l'infini et voyons ce qu'il advient des généalogies associées. Comme mentionné au début de ce paragraphe, la proposition suivante correspond à la Proposition 1 de [**Sch00**], mais le résultat original (utilisant une autre caractérisation des Ξ-coalescents) constitue la première partie du Théorème 2.1 de [**MS01**]. Notons $R^{[k,N]}$ le processus généalogique des individus $1,\dots,k$ vivant au temps $t=0$, et ν_N le vecteur aléatoire donnant la répartition des descendants dans un modèle de Cannings dont la taille totale de la population est N.

PROPOSITION 1.31. *Supposons que pour tout* $r > 0$ *et* $k_1, \ldots, k_r \geq 2$, *les limites*

$$\lim_{N\to\infty} \frac{\mathbb{E}[(\nu_{1,N})_{k_1} \cdots (\nu_{r,N})_{k_r}]}{N^{k_1+\ldots+k_r-r} c_N}$$

existent et que $c_N \to 0$ *lorsque* $N \to \infty$. *Alors il existe un coalescent à collisions multiples et simultanées* R *tel que pour tout* $k \in \mathbb{N}$, *la suite* $\big(\{R^{[k,N]}_{\lfloor t/c_N \rfloor},\ t \geq 0\}\big)_{N\in\mathbb{N}}$ *de processus markoviens à temps continu et à valeurs dans* $\mathcal{P}_k$ *converge pour la topologie de Skorokhod vers* $R\big|_{\mathcal{P}_k}$. *La mesure* Ξ *associée à* R *est caractérisée par*

$$\lambda_{b;k_1,\ldots,k_r;0} = \lim_{N\to\infty} \frac{\mathbb{E}[(\nu_{1,N})_{k_1} \cdots (\nu_{r,N})_{k_r}]}{N^{k_1+\ldots+k_r-r} c_N},$$

pour tous $r \geq 1$, $k_1, \ldots, k_r \geq 2$ *et* $b = \sum_{i=1}^r k_i$.

Ce résultat contient beaucoup d'informations. Premièrement, la bonne échelle de temps à considérer est c_N^{-1}, que l'on avait obtenu auparavant comme nombre moyen de générations à remonter pour trouver le premier ancêtre commun à deux individus. Deuxièmement, la propriété de cohérence est conservée à la limite, puisque le processus généalogique que l'on obtient pour chaque k est la restriction à $\mathcal{P}_k$ d'un même Ξ-coalescent. Troisièmement, en utilisant l'équivalence asymptotique $(N)_n \sim N^n$ lorsque $N \to \infty$ et n est fixé, on reconnaît les coefficients p_N dans les critères de convergence de la Proposition 1.31, qui se réécrivent alors

$$\lambda_{b;k_1,\ldots,k_r;0} = \lim_{N\to\infty} \frac{p_N(k_1,\ldots,k_r)}{c_N} \ (\leq 1).$$

Ainsi, puisque le nombre de transitions possibles est fini lorsque k est fixé, si $p_N(k_1,\ldots,k_r) \ll c_N$ dès que $r \geq 2$ ou $k_1 \geq 3$, chaque paire de blocs de $R^{[k,N]}$ fusionne bien plus rapidement qu'elle n'est prise dans un événement impliquant plus de blocs et le coalescent limite n'évolue que par fusions de deux blocs à la fois. Au contraire, si $p_N(k_1,\ldots,k_r)$ et c_N sont asymptotiquement comparables, les événements de coalescence impliquant plus de deux blocs arrivent sur la même échelle de temps que les fusions de paires de blocs et à la limite ces deux types de collisions sont présents. En considérant $p_N(3)$ et $p_N(2,2)$, on voit alors de quelle manière apparaissent les collisions multiples et simultanées. En effet,

$$\frac{p_N(3)}{c_N} = \frac{\mathbb{E}[\nu_1(\nu_1-1)(\nu_1-2)]}{(N-2)\mathbb{E}[\nu_1(\nu_1-1)]} \qquad \text{et} \qquad \frac{p_N(2,2)}{c_N} \sim \frac{\mathbb{E}[\nu_1(\nu_1-1)\nu_2(\nu_2-1)]}{N\mathbb{E}[\nu_1(\nu_1-1)]}.$$

De manière heuristique, si la probabilité qu'un seul individu produise un nombre assez important de descendants (qui se traduit dans les moments de ν_1) décroît trop rapidement lorsque $N \to \infty$, alors celle que l'individu 3 ait aussi été choisi parmi la même famille que 1 et 2 est négligeable et $p_N(3)/c_N \to 0$: *la seule manière d'obtenir des coalescences multiples à la limite est donc d'autoriser certains individus à produire parfois un nombre « conséquent » d'individus de la nouvelle génération* (puisque le nombre total de descendants est toujours égal à N, cela implique que de nombreux autres individus ne se reproduisent pas). Ensuite, si la probabilité que deux individus produisent simultanément assez de descendants (qui se lit dans les moments de (ν_1, ν_2)) décroît trop rapidement, il devient difficile d'échantillonner uniformément au hasard deux individus dans chacune des deux familles correspondantes, puisqu'au moins l'une d'entre elle est « petite ». C'est cette fois $p_N(2,2)/c_N$ qui tend vers 0 : *pour que le coalescent limite ait des collisions simultanées, il faut donc que plusieurs individus puissent chacun être responsables d'un nombre « assez grand »* *d'individus de la nouvelle génération.* Si aucune de ces deux conditions n'est remplie, on obtient un coalescent de Kingman, avec des fusions uniquement binaires.

REMARQUE 1.32. *Le fait que les coefficients $\lambda_{b;k_1,\dots,k_r;0}$ caractérisent la mesure Ξ n'est pas un résultat de la Proposition 1 de* [**Sch00**]*, mais découle d'une formule de récurrence donnée par la cohérence du Ξ-coalescent et permettant d'obtenir tous les coefficients $\lambda_{b;k_1,\dots,k_r;s}$ à partir de ceux où $s = 0$ (cf.* [**Sch00**]*, p.4).*

La Proposition 7 de [**Sch00**] montre en outre que tous les Ξ-coalescents peuvent être obtenus comme limite de généalogies de modèles de Cannings. On trouvera dans [**EW06**] un exemple de convergence vers des Λ-coalescents à partir de modèles de Cannings dans lesquels à chaque génération, un seul individu produit un nombre potentiellement grand de descendants tandis que les autres ont 0 ou 1 descendant (suivi d'une étude des conséquences sur la diversité génétique de telles populations). Dans [**Sch03**], Schweinsberg considère un modèle de Cannings construit à partir d'un processus de Galton-Watson surcritique dont la queue de distribution du nombre de descendants décroît en k^{-a} et obtient lorsque la taille de la population tend vers l'infini un coalescent dont la forme (i.e., l'existence de collisions multiples ou simultanées) dépend de a.

Remarquons que les populations dont on considère la généalogie ici ont relativement peu d'intérêt : contrairement aux processus de Galton-Watson, on force leur taille à rester constante et donc si on ne différencie pas les individus, rien ne se passe. Par ailleurs, un modèle de Cannings dont la taille de la population est infinie n'a pour le moment pas grand sens. Introduisons donc la notion de *type* (ou d'*allèle* pour des problèmes de génétique) comme étant une caractéristique portée par un individu et transmise de parent à descendant. Puisque ce sont les généalogies qui nous intéressent principalement, on suppose que la transmission est parfaite (pas de *mutation*) et pour garder des processus échangeables, on fait l'hypothèse supplémentaire que le type d'un individu n'a aucune influence sur sa capacité à se reproduire (aucun allèle n'a d'*avantage sélectif*). On supposera toujours que l'espace des allèles possibles est compact, et la plupart du temps il ne contiendra que deux éléments a et A. Les questions auxquelles on s'intéresse alors sont l'évolution des sous-populations d'un type particulier, la possibilité de leur extinction, ou encore celle de *fixation* d'un allèle, i.e., la probabilité qu'il envahisse complètement la population. Ces questions ne sont évidemment pas sans lien avec la généalogie sous-jacente. Les différentes relations de dualité que nous verrons dans les paragraphes suivants permettent en fait d'exprimer des quantités comme l'*homozygocité* au temps $t > 0$ (i.e., la probabilité que deux individus pris au hasard dans la population au temps t soient du même type) en fonction de la probabilité que ces individus aient un ancêtre commun au temps 0.

2.1.3. *Dualité.* La taille des populations considérées dans le paragraphe précédent est finie pour chaque $N \in \mathbb{N}$; il est donc possible de donner un sens à la généalogie de k individus choisis au hasard et de relier celle-ci à l'évolution de la population dans le sens *normal* du temps (i.e., vers les temps positifs). Après le passage à la limite $N \to \infty$ en revanche, échantillonner uniformément au hasard k individus dans une population de taille infinie devient problématique et l'existence et les mécanismes d'évolution de processus généalogiques sous-jacents sont encore moins clairs. Dans certains cas tels que les modèles de Cannings, on peut construire une représentation graphique (plus riche que le processus initial) qui permet de définir un « individu » d'une population infinie et qui contient assez d'information pour tracer à la fois l'évolution vers le futur de la population et celle vers le passé des relations généalogiques entre les individus. C'est l'une des raisons d'être des constructions de type *look-down* introduites dans [**DK99**], qui permettent par exemple à Birkner & al. dans [**B&al05**] de plonger certains Λ-coalescents dans des modèles de CSBP de mécanisme de branchement stable. On obtient alors une *dualité forte* entre le modèle de population et les processus généalogiques associés.

Une version plus faible mais néanmoins utile est celle de *dualité fonctionnelle* ou *faible*, qui permet de relier les marginales de dimension finie de deux processus pour lesquels « le temps n'évolue pas dans le même sens ». La définition suivante est tirée de [**Eth00**], mais on en trouvera une version plus générale dans le chapitre 4.4 de [**EK86**]. Pour rappel, on note $D_E[0,\infty)$ l'espace des fonctions càdlàg sur $[0,\infty)$ à valeurs dans E.

DÉFINITION ET THÉORÈME 1.33. *Soient E_1, E_2 deux espaces métriques munis de leurs tribus boréliennes, et X (resp., Y) un processus markovien à trajectoires dans $D_{E_1}[0,\infty)$ (resp., $D_{E_2}[0,\infty)$). Soient f, g des fonctions bornées sur $E_1 \times E_2$ vérifiant les conditions suivantes :*

1. *Pour tout $y \in E_2$, $f(\cdot, y)$ et $g(\cdot, y)$ sont des fonctions continues sur E_1.*
2. *Pour tout $x \in E_1$, $f(x, \cdot)$ et $g(x, \cdot)$ sont des fonctions continues sur E_2.*
3. *Pour tout $y \in E_2$,*
$$f(X_t, y) - \int_0^t g(X_s, y)\, ds$$
est une martingale pour la filtration canonique associée à X.
4. *Pour tout $x \in E_1$,*
$$f(x, Y_t) - \int_0^t g(x, Y_s)\, ds$$
est une martingale pour la filtration canonique associée à Y.

Alors, pour tout $t \geq 0$ on a
$$\mathbb{E}^X\big[f(X_t, Y_0)\big] = \mathbb{E}^Y\big[f(X_0, Y_t)\big]$$
et les processus X et Y sont dits duaux pour la fonction f.

Lorsque la famille de fonctions f pour lesquelles X et Y sont duaux est assez grande, tout autre processus dual de X pour les mêmes fonctions possède les mêmes (en loi) marginales de dimension finie que Y, ce qui permet par exemple de montrer l'unicité de la solution d'un problème de martingales (cf. chapitre 4.4 de [**EK86**]).

Remarquons en exemple que nous avons déjà croisé des processus duaux dans le paragraphe 1.2.1. En effet, le super-mouvement brownien X possède un dual déterministe $Y = \{u(t, \cdot),\ t \geq 0\}$ à valeurs dans un ensemble de fonctions : on a vu dans la Définition 1.6 que pour tout $f \in \mathcal{B}_b^+(\mathbb{R}^d)$,
$$\mathbb{E}_\mu\big[e^{-\langle X_t, f\rangle}\big] = e^{-\langle \mu, u(t,\cdot)\rangle},$$
où $u(t,x)$ est solution de l'équation intégrale (1.13). En particulier, $u(0,\cdot) = f$. Des propriétés de u découlent alors certains résultats sur les marginales de dimension finie du super-mouvement brownien.

Pour revenir aux processus généalogiques, on verra dans le paragraphe 2.2.1 que le modèle *stepping stone* de population structurée en des sous-populations discrètes (ou îles) a pour dual un système de marches aléatoires coalescentes, dont le mouvement entre les îles est dicté par le mécanisme de migration des individus dans le sens *normal* du temps et les coalescences au sein d'une sous-population correspondent au mécanisme de reproduction dans l'île correspondante. Ainsi, le dual faible que l'on obtient peut être vu comme le processus généalogique d'un nombre fini d'individus échantillonnés dans la population totale. Dans le paragraphe 2.2.2, la dualité faible avec un système de processus de Lévy coalescents sera même la clé de l'existence du processus Λ-Fleming-Viot spatial, dont on pourra à nouveau interpréter le dual comme donnant la généalogie d'un nombre fini d'individus choisis dans la population. Soulignons cependant que qualifier de *généalogique* un processus obtenu par dualité faible reste du domaine de l'interprétation : seule une dualité forte par une construction de type *look-down* ou graphique (comme pour le modèle du votant) garantit que la généalogie d'un échantillon d'individus dans un modèle de population donné évolue réellement selon le mécanisme dual prescrit. On consultera [**Tay09**] pour un exemple dans lequel plusieurs modèles de population convergent vers la même limite tandis que leurs processus ancestraux ont des limites distinctes.

2.2. Modèles de populations structurées et généalogies associées. Intéressons-nous à présent à des populations réparties sur un espace.

2.2.1. *Espace géographique discret.* Une première manière d'introduire une structure spatiale est de supposer que la population est découpée en communautés discrètes (aussi appelées *îles*, même si la structure à laquelle on s'intéresse n'est pas nécessairement géographique), situées aux sommets d'un graphe. Dans le cas du graphe complet, dans lequel tous les sommets sont à même distance les uns des autres, on obtient alors le *modèle d'îles* de Wright introduit dans [**Wri31**]. Le modèle *stepping stone* introduit par Kimura dans [**Kim53**] est quant à lui défini sur le graphe $\mathbb{Z}^d$ (on prendra $d = 2$ pour une structure géographique raisonnable). Dans ce qui suit, l'espace des types est $\{a, A\}$ et $p_i(t)$ donne la fréquence au temps $t \geq 0$ des individus de type a dans la communauté indicée par i (la fréquence des individus de type A est donc $1 - p_i(t)$).

Modèle d'îles de Wright

C'est le modèle de structure le plus simple que l'on puisse imaginer. La population y est répartie en D îles ($D \in \mathbb{N} \cup \{\infty\}$), toutes connectées entre elles. Dans chaque communauté i, la fréquence p_i évolue grâce à la dynamique de reproduction interne à la communauté (identique d'une île à l'autre), mais est aussi modifiée par le flux de migrants provenant des autres îles. Pour écrire de manière un peu formelle (mais qui comprend le cas où le nombre d'îles est infini et celui où la taille d'une sous-population et donc le nombre de migrants sont infinis) que chaque migrant provient de n'importe quelle île avec la même probabilité, considérons le reste de la population comme un réservoir auquel chaque île $j \neq i$ contribue de manière identique. La probabilité qu'un arrivant sur l'île i soit de type a est simplement la fréquence de a dans ce réservoir.

Ce modèle de base comporte beaucoup de symétries, que l'on peut briser par exemple en supposant que les tailles des sous-populations sont différentes ou fluctuent au cours du temps. On appellera cependant *modèle d'îles de Wright* tout modèle dans lequel le graphe sous-jacent est complet. Ce cadre simple nous permettra dans le chapitre 4 de mettre l'accent sur les mécanismes conduisant à l'existence de coalescences multiples et simultanées dans la généalogie d'un échantillon d'individus pris dans la population totale, lorsque le nombre d'îles est très grand.

Modèle stepping stone

Afin d'instaurer une notion de distance entre les îles, plaçons nous à présent sur $\mathbb{Z}^d$. A nouveau, les individus se reproduisent à l'intérieur de leur communauté et des migrations d'individus d'une île à l'autre introduisent des dépendances entre les fréquences locales d'allèle a.

DÉFINITION 1.34. *Dans le modèle stepping stone, les fréquences p_i évoluent selon le système d'équations différentielles stochastiques suivant : pour tout $i \in \mathbb{Z}^d$,*

$$dp_i = \sum_{j \neq i} m_{ji}(p_j - p_i) + \sqrt{\frac{1}{N_e}\, p_i(1 - p_i)}\, dB_i,$$

où $\{B_i,\ i \in \mathbb{Z}^d\}$ *est une collection de mouvements browniens standards indépendants et les coefficients* m_{ji} *satisfont pour tout* i

$$\sum_{j\neq i} m_{ij} = \sum_{j \neq i} m_{ji}. \tag{1.28}$$

Dans cette définition, chaque île contient moralement une infinité d'individus et la dynamique de reproduction intra-île suit une diffusion de *Wright-Fisher*, donnée par les équations différentielles stochastiques indépendantes

$$dp_i = \sqrt{\frac{1}{N_e}\, p_i(1-p_i)}\, dB_i,$$

où l'on reconnaît un terme de compétition $p_i(1-p_i)$ entre les deux allèles et N_e est un paramètre contrôlant la vitesse de la diffusion. Les fréquences locales sont reliées entre elles par les termes de migration, dans lesquels m_{ji} représente le taux auquel la sous-population j envoie des migrants sur l'île i. Les relations (1.28) garantissent un équilibre entre les flux sortants et entrants pour chaque communauté.

Cette fois, la structure du graphe apparaît vraiment puisque les paramètres m_{ij} induisent une sorte de distance entre les communautés : si par exemple $m_{ij} = 0$ dès lors que $\|i-j\| > R$, un migrant ne peut descendre que d'un individu d'une communauté située dans un rayon R autour de son île d'adoption et le temps nécessaire pour remonter au premier ancêtre commun de deux individus appartenant à des îles très éloignées sera a priori beaucoup plus long que si ces individus avaient été choisis dans des îles proches (en supposant qu'un tel ancêtre existe).

La définition d'une généalogie sous-jacente et les relations qu'elle induit entre différentes îles se reformulent agréablement grâce à un argument de dualité faible (cf. §2.1.3). En effet, on peut montrer que le modèle stepping stone est dual du système de marches aléatoires coalescentes $(\{n_i(t),\ i \in \mathbb{Z}^d\})_{t\geq 0}$, dans lequel $n_i(t)$ est le nombre de particules appartenant à la communauté i au temps t et l'évolution du système est donnée par :

- pour tout $i \in \mathbb{Z}^d$, $n_i \mapsto n_i - 1$ à taux $\frac{1}{N_e}\binom{n_i}{2}$;
- pour toute paire $i \neq j$, $(n_i, n_j) \mapsto (n_i - 1, n_j + 1)$ à taux $n_i m_{ji}$.

En utilisant la Définition 1.33, on peut vérifier que pour tout $t \geq 0$

$$\mathbb{E}\Big[\prod_{i\in\mathbb{Z}^d} p_i(t)^{n_i(0)}\Big] = \mathbb{E}\Big[\prod_{i\in\mathbb{Z}^d} p_i(0)^{n_i(t)}\Big]. \tag{1.29}$$

Le système $\{n_i,\ i \in \mathbb{Z}^d\}$ correspond bien à l'idée que l'on se fait d'un processus généalogique potentiel. En effet, on peut obtenir la diffusion de Wright-Fisher comme dans le paragraphe 2.1.2 à partir d'un modèle de Cannings à deux types, dont la généalogie converge vers un coalescent de Kingman (changé de temps par un facteur $1/N_e$). On s'attend donc à ce que les lignages ancestraux fusionnent uniquement par paires et que le taux de coalescence de chacune des $\binom{n_i}{2}$ paires dans la communauté i soit $1/N_e$. D'autre part, puisque m_{ji} est le taux auquel un individu de l'île j envoie un descendant sur l'île i, en renversant le temps on obtient que m_{ji} est également le taux auquel chaque lignage ancestral appartenant à l'île i migre vers l'île j. Pour ce qui est des relations entre les fréquences p_i sur différentes îles, grâce à (1.29) on voit par exemple que le moment d'ordre k de $p_i(t)$ dépend de la généalogie de k individus échantillonnés dans l'île i, et que la covariance des fréquences d'allèle a aux sites i et j de $\mathbb{Z}^d$ s'exprime en termes des fréquences initiales et de la probabilité que deux marches aléatoires partant des sites i et j respectivement fusionnent avant le temps t en une seule particule (autrement dit, que deux lignages ancestraux partant des îles i et j atteignent leur premier ancêtre commun avant le temps t dans le passé).

Puisque dans chacun de ces modèles les lignages ancestraux ne peuvent fusionner qu'au sein d'une même communauté, il n'est plus possible de décrire la généalogie de k individus par un coalescent échangeable : il faut également garder en mémoire l'indice de l'île dans laquelle se trouve chaque ancêtre et les migrations font partie intégrante du processus généalogique. La manière la plus répandue de représenter une généalogie structurée est de considérer la collection $(\{n_i(t),\ i \in I\})_{t\geq 0}$ donnant le nombre d'ancêtres de l'échantillon dans chaque île comme on l'a fait ci-dessus, I étant l'ensemble des indices des communautés. On peut aussi souhaiter conserver plus d'informations et définir un processus sur l'ensemble des partitions marquées de $[k]$. L'étiquette de chaque bloc au temps t correspond alors à l'indice de l'île dans laquelle se trouve l'ancêtre correspondant à cet instant. On pourra consulter [**Not90**] et [**WH98**] pour une étude de coalescents de Kingman structurés et de différentes statistiques de la diversité génétique associées.

2.2.2. *Espace géographique continu.* Les modèles présentés dans le paragraphe précédent ne rendent pas toujours bien compte de l'évolution d'une population réelle. En effet, la plupart des populations évoluent dans un espace continu et il n'est pas toujours évident de les découper en des communautés discrètes de manière non-artificielle. Pour pallier ce défaut, une idée naturelle est de partir d'un modèle stepping stone, dans lequel $m_{ij} = m\mathbb{I}_{\{\|i-j\|=1\}}$ par exemple, et de passer à une limite diffusive en renormalisant la distance entre deux sites voisins par ε et le temps par ε^{-2} comme on l'a fait pour obtenir des superprocessus dans le paragraphe 1.2.1 (remarquons que l'on travaille ici avec des fréquences $p_i \in [0,1]$; il n'y a donc pas besoin de renormaliser la masse de chaque île). En dimension 1, on obtient alors un processus limite $(\{p(x,t),\ x \in \mathbb{R}\})_{t\geq 0}$, où $p(x,t)$ est la fréquence des individus de type a au site x et à l'instant t et satisfait l'équation aux dérivées partielles stochastique

$$dp = \frac{m}{2}\Delta p\, dt + \sqrt{\frac{1}{N_e}p(1-p)}\, dW, \tag{1.30}$$

où W est un bruit blanc sur $\mathbb{R}\times\mathbb{R}_+$. Ce résultat est dû à Shiga (cf. [**Shi88**]), qui montre également que le processus généalogique dual converge en loi vers un système de mouvements browniens coalescents dans lequel chaque paire de lignages fusionne à un taux dépendant du temps local en 0 de leur différence. On voit alors se profiler les problèmes potentiels en dimension supérieure : deux mouvements browniens ne se rencontrant jamais, deux lignages n'ont aucune chance d'avoir un ancêtre commun. En fait, on peut montrer que (1.30) n'a pas de solution en dimension supérieure ou égale à deux et le modèle discret renormalisé converge simplement vers l'équation de la chaleur. Cette approche ne donne donc pas de résultat convenable en dimension $d \geq 2$.

En réalité, notre liste d'exigences est assez longue : on souhaite arriver à un modèle dans lequel

- la *population locale effective* reste bornée, pour tenir compte des phénomènes de régulation locale du nombre d'individus. En d'autres termes, on veut garder la signature des populations finies en autorisant des coalescences multiples voire simultanées dans la généalogie d'un échantillon d'individus « géographiquement proches » ;
- les reproductions se font de manière locale, pour modéliser le fait que les descendants ont plus de chance de naître dans un voisinage de leur parent ;
- les décès sont aussi corrélés géographiquement, suivant l'idée que si un événement catastrophique (orage, sécheresse, tremblement de terre, ...) survient, il est très probable qu'il affecte les individus d'une zone donnée de l'espace ;
- on peut définir un processus généalogique cohérent au sens de la Propriété 1.27, qui semble indispensable pour un processus qui retrace la généalogie de k individus.

Ces différents objectifs sont remplis avec le *processus* Λ*-Fleming-Viot spatial*, introduit par Barton et Etheridge et dont la première description se trouve dans [**Eth08**]. Si on note K l'espace compact des allèles possibles ($K = [0,1]$ ou $\{A, a\}$ par exemple), la population à l'instant t est alors représentée par une fonction $x \in \mathbb{R}^d \mapsto \rho(t,x,\cdot) \in \mathcal{M}_1(K)$ (l'ensemble des mesures de probabilité sur K). La définition du processus est la suivante :

DÉFINITION 1.35. *Soit μ une mesure σ-finie sur $(0,\infty)$, $\{\nu_r,\ r > 0\}$ une collection de mesures de probabilité sur $[0,1]$ et Π un processus ponctuel de Poisson sur $\mathbb{R}_+ \times \mathbb{R}^d \times (0,\infty)$ d'intensité $dt \otimes dx \otimes \mu(dr)$. Le* processus Λ-Fleming-Viot spatial *est le processus markovien $(\{\rho(t,x,\cdot),\ x \in \mathbb{R}^d\})_{t\geq 0}$ à valeurs dans l'ensemble des fonctions $\varrho : \mathbb{R}^d \to \mathcal{M}_1(K)$, qui évolue de la manière suivante : si $(t,x,r) \in \Pi$, on choisit $u \in [0,1]$ suivant la loi $\nu_r(du)$, un point z uniformément au hasard dans la boule $B(x,r)$ et un type k suivant la loi $\rho(t-,z,\cdot)$. Pour tout $y \in B(x,r)$, la nouvelle valeur de ρ est donnée par*

$$\rho(t,y,\cdot) = (1-u)\rho(t-,y,\cdot) + u\delta_k.$$

Les sites en dehors de $B(x,r)$ ne sont pas affectés, i.e., $\rho(t,y,\cdot) = \rho(t-,y,\cdot)$ pour tout $y \notin B(x,r)$.

L'interprétation que l'on peut faire de ce modèle est que chaque site contient une infinité d'individus et $\rho(t,x,\cdot)$ est la distribution du type d'un individu que l'on échantillonnerait au temps t au site x. Lorsqu'un événement survient, une fraction u (dépendant de la taille de la zone affectée) des individus dans $B(x,r)$ est tuée et remplacée par les descendants d'un individu choisi uniformément au hasard parmi la population contenue dans $B(x,r)$ juste avant l'événement. L'intérêt de choisir cet individu en deux étapes (au lieu de choisir directement un type selon la distribution des types moyennée sur $B(x,r)$) est que l'on garde ainsi en mémoire sa position spatiale, ce qui nous servira pour retracer l'évolution d'un lignage ancestral.

REMARQUE 1.36. (Processus Λ-Fleming-Viot non-spatial) *Une version non-spatiale de ce processus a été introduite par Bertoin et Le Gall dans* [**BL03**]. *La population au temps t est représentée par une mesure $\tilde{\rho}(t,\cdot) \in \mathcal{M}_1(K)$, dont l'évolution dépend d'une mesure finie Λ sur $[0,1]$. Dans le cas où $\Lambda(\{0\}) = 0$, cette évolution peut être décrite en termes d'un processus ponctuel de Poisson $\tilde{\Pi}$ sur $\mathbb{R}_+ \times [0,1]$ d'intensité $dt \otimes u^{-2}\Lambda(du)$: si $(t,u) \in \tilde{\Pi}$, un type k est choisi selon $\tilde{\rho}(t-,\cdot)$ et une fraction u de la population est tuée et remplacée par les descendants de l'individu choisi :*

$$\tilde{\rho}(t,\cdot) = (1-u)\tilde{\rho}(t-,\cdot) + u\delta_k.$$

Cette description rappelle la construction poissonienne des coalescents donnée dans la Remarque 1.28 et effectivement, le processus Λ-Fleming-Viot est dual du Λ-coalescent de même mesure. Par analogie avec les superprocessus, on peut obtenir les processus Λ-Fleming-Viot comme limites de modèles de Cannings avec types (les mutations correspondant au déplacement spatial des superprocessus) renormalisés pour avoir une masse totale de population constante égale à 1.

REMARQUE 1.37. *Le processus Λ-Fleming-Viot spatial peut être obtenu comme limite d'un processus dans lequel chaque région de l'espace ne contient qu'un nombre fini d'individus, lorsque la densité d'individus tend vers l'infini. Ce modèle à taille de population localement finie est également introduit dans* [**Eth08**] *et les questions d'ergodicité et de survie ou extinction locale sont l'objet de* [**BEH09**].

L'existence du processus Λ-Fleming-Viot spatial est obtenue dans le paragraphe 4 du chapitre 5, grâce à la méthode utilisée par Evans dans [**Eva97**] pour construire des modèles stepping stone en espace continu comme duaux de processus markoviens coalescents (que l'on peut ensuite interpréter comme la généalogie d'individus du modèle de population ainsi construit). Dans

[**Eva97**], chaque lignage se déplace indépendamment des autres selon la loi d'un processus markovien ξ et lorsque plusieurs lignages se rencontrent, ils fusionnent instantanément en un unique lignage qui reprend alors son déplacement spatial selon la loi de ξ jusqu'à sa prochaine rencontre avec d'autres lignages, etc. Voyons donc à quoi ressemble la généalogie que l'on pourrait associer au processus Λ-Fleming-Viot spatial et quelles sont les conditions pour que celle-ci, partant d'un nombre fini d'individus, corresponde à un processus de Markov bien défini.

Considérons un seul lignage. Puisque les individus dans le sens *normal* du temps ne se déplacent pas, sa position spatiale ne change que lorsqu'il est créé au cours d'un événement de Π (le processus ponctuel de Poisson de la Définition 1.35), i.e., lorsqu'il fait partie de la fraction u d'individus nés du parent choisi. Le lignage adopte alors la position de ce parent, qui est par construction uniformément distribuée sur la boule dans laquelle l'événement se produit. Autrement dit, pour qu'un lignage actuellement en 0 saute en $x \in \mathbb{R}^d$, il faut que survienne un événement de paramètres (z, r, u) tels que

- 0 et x appartiennent à $B(z, r)$, soit $r \geq \|x\|/2$ et $z \in B(0, r) \cap B(x, r)$, ce qui arrive à taux $\mathrm{Vol}\big(B(0, r) \cap B(x, r)\big) := L_r(x)$;
- le lignage que l'on suit fasse partie des nouveaux individus créés au site 0, ce qui arrive avec probabilité u ;
- le parent choisi est situé en x, ce qui arrive avec probabilité $(c_d r^d)^{-1}dx$ si on note c_d le volume de la boule unité en dimension d.

Par invariance par translation spatiale et temporelle de la loi des événements, on obtient alors que, quelle que soit la position ξ du lignage et le temps auquel on le considère, l'intensité à laquelle il saute en $\xi + x$ est donnée par la mesure

$$J(dx) := \bigg(\int_{\|x\|/2}^{\infty} \int_0^1 \frac{L_r(x)}{c_d r^d}\, u\nu_r(du)\mu(dr) \bigg) dx. \tag{1.31}$$

Le déplacement spatial (supposé) d'un lignage est donc un processus de sauts dont les incréments sont stationnaires et indépendants. Pour en faire un processus de Lévy sur $\mathbb{R}^d$, il nous faut imposer que

$$\int_{\mathbb{R}^d} (1 \wedge \|x\|^2) J(dx) < \infty. \tag{1.32}$$

Passons à deux lignages. Leurs déplacements sont corrélés, puisqu'ils utilisent potentiellement les mêmes événements de Π pour se déplacer. Par ailleurs, pour que deux lignages fusionnent il faut qu'ils soient tous deux créés au cours d'un même événement de Π. Si $x \in \mathbb{R}^d$ est la séparation actuelle des deux lignages, le taux auquel ils fusionnent est donc

$$\int_{\|x\|/2}^{\infty} \int_0^1 L_r(x) u^2 \nu_r(du)\mu(dr), \tag{1.33}$$

où $L_r(x)$ est le volume de la région des centres possibles pour qu'un événement de taille r englobe les deux lignages et u^2 est la probabilité que ceux-ci soient alors contenus dans la fraction u de nouveaux-nés dans leurs sites respectifs. Si seulement l'un d'entre eux est affecté par un événement qui les englobe géographiquement tous les deux (ce qui arrive avec probabilité $2u(1-u)$), l'un saute sur le site contenant son parent et l'autre ne bouge pas. Le raisonnement est identique pour plus de deux lignages, même si écrire explicitement le taux auquel certains d'entre eux sautent et fusionnent tandis que les autres ne bougent pas devient de plus en plus complexe. Notons $\mathcal{A}$ le processus généalogique (supposé) d'un nombre fini d'individus, dont on précisera l'espace d'états un peu plus tard.

Comme pour les Λ-coalescents non-spatiaux, on impose que le taux auquel deux lignages fusionnent soit fini. Puisqu'il est maximal lorsque les deux lignages appartiennent au même site, ce que l'on peut vérifier dans (1.33), on demande donc que

$$\int_0^\infty \int_0^1 r^d u^2 \nu_r(du)\mu(dr) < \infty. \tag{1.34}$$

En fait, cette condition n'est pas suffisante : la méthode d'Evans permet bien de construire un modèle stepping stone en espace continu dual de $\mathcal{A}$ sous les hypothèses (1.32) et (1.34), mais si on autorise le déplacement d'un lignage à accumuler de très petits sauts, le générateur de $\mathcal{A}$ ne peut pas s'écrire sous la forme d'une intégrale des incréments contre l'intensité des sauts (phénomène comparable à la correction impliquant la dérivée première dans le générateur d'un seul processus de Lévy). C'est cette écriture qui est la clé de l'identification du dual de $\mathcal{A}$ au processus que l'on a décrit en fonction de Π (voir le paragraphe 4.2 du chapitre 5). On doit donc imposer la condition plus forte

$$\int_0^\infty \int_0^1 r^d u \, \nu_r(du)\mu(dr) < \infty, \tag{1.35}$$

sous laquelle chaque lignage saute à taux fini, ce qui implique en outre que (1.32) est satisfaite et deux lignages fusionnent également à taux fini. La formulation sous forme de processus ponctuel d'événements au cours desquels les lignages décident indépendamment de leur sort rend le processus généalogique $\mathcal{A}$ cohérent et donc le taux auquel plusieurs lignages fusionnent est borné par celui auquel deux d'entre eux coalescent. La généalogie que l'on obtient pour un échantillon fini est ainsi un processus de sauts à taux fini, markovien et cohérent dans lequel, à la différence du cadre de travail d'Evans, plusieurs lignages ont des déplacements corrélés et ne fusionnent pas en se rencontrant mais en étant affectés par le même événement de Π.

REMARQUE 1.38. *Le problème conduisant à la condition (1.35) est plus profond que l'incapacité à identifier le dual des processus de Lévy coalescents au processus auquel on souhaite arriver. Il vient du fait que, contrairement à un processus Λ-Fleming-Viot non-spatial, à un site y fixé le choix du type de l'individu qui se reproduit ne dépend pas de $\rho(t-,y,\cdot)$, mais de la distribution des allèles dans toute la boule où a lieu l'événement. Les sauts de très petite amplitude de $\rho(t,y,\cdot)$ ne se compensent donc pas en moyenne et on doit imposer que le terme de dérive du générateur infinitésimal de ρ à un site donné soit fini : en considérant le pire scénario, on obtient la condition (1.35). On renvoie au paragraphe précédant l'introduction de l'Hypothèse 5.5 dans le chapitre 5 pour un calcul explicite du générateur du processus ρ agissant sur des fonctions de la distribution des types à un site donné.*

Décrivons à présent la relation de dualité entre ρ et $\mathcal{A}$. Tout d'abord, on représente les relations généalogiques entre n individus au temps t dans le passé par une partition marquée de $[n]$ dans laquelle chaque bloc contient les étiquettes des descendants d'un même individu vivant au temps $-t$, dont la position spatiale est donnée par la marque du bloc. Notons $\mathcal{P}_n^\ell$ l'ensemble des partitions marquées de $[n]$ et $\wp_n(x_1,\ldots,x_n)$ la partition marquée triviale $\{(\{1\},x_1),\ldots,(\{n\},x_n)\}$. Appelons Ξ l'espace des fonctions mesurables de $\mathbb{R}^d$ dans $\mathcal{M}_1(K)$ sur lequel on travaille (une définition précise et la topologie associée sont données dans le paragraphe 4.1) du chapitre 5. Les fonctions auxquelles on s'intéresse sont de la forme

$$I_n(\varrho;\Phi) := \int_{(\mathbb{R}^d)^n} \left\langle \bigotimes_{1\le i\le n} \varrho(x_i), \Phi(x_1,\ldots,x_n)\right\rangle dx_1\ldots dx_n,$$

où $\varrho \in \Xi$ et $\Phi \in L^1(C(K^n))$, l'espace des fonctions mesurables de $(\mathbb{R}^d)^n$ dans $C(K^n)$ telles que $\int \|\Phi(x)\|dx < \infty$. Définissons également pour tout $\varrho \in \Xi$, $\pi = \{(a_1, x_1), \ldots, (a_k, x_k)\} \in \mathcal{P}_n^\ell$ et $F : K^n \to \mathbb{R}$ mesurable bornée

$$\Upsilon_n(\varrho; \pi; F) := \int_{K^k} F(v_{a^{-1}(1)}, \ldots, v_{a^{-1}(n)})\varrho(x_1)(dv_1) \ldots \varrho(x_k)(dv_k),$$

où $a^{-1}(i)$ est l'indice de l'unique bloc de π contenant i. En d'autres termes, la même variable v_i est utilisée pour toutes les coordonnées appartenant au bloc a_i. Malgré les différences avec les hypothèses d'Evans que l'on a pu constater, la méthode employée dans [**Eva97**] reste valable pour construire le processus Λ-Fleming-Viot spatial et le Théorème 5.18 assure alors qu'il existe un unique processus markovien ρ, satisfaisant la propriété de Feller, tel que pour tout $n \in \mathbb{N}$, $\Phi \in L^1(C(K^n))$ et $\varrho \in \Xi$ on a

$$\mathbb{E}_\varrho\big[I_n(\rho_t, \Phi)\big] = \int_{(\mathbb{R}^d)^n} \mathbb{E}_{\wp_n(\mathbf{x})}\big[\Upsilon_n\big(\varrho; \mathcal{A}_t; \Phi(x_1, \ldots, x_n)\big)\big]dx_1 \ldots dx_n,$$

où $\mathbf{x} = (x_1, \ldots, x_n)$. On identifie ensuite le processus ρ obtenu par dualité à celui dont l'évolution est donnée par le processus ponctuel d'événements.

Remarquons que cette forme de dualité est analogue à celle que l'on obtient pour le modèle stepping stone en espace discret. Comme dans (1.29), les relations entre les distributions d'allèles au temps t à n sites donnés s'écrivent en fonction de la généalogie de n particules entre les instants t et 0 et de la valeur initiale ϱ : si plusieurs individus échantillonnés au temps t ont déjà atteint leur ancêtre commun au temps 0, alors leurs allèles sont nécessairement identiques et distribués selon la valeur de ϱ au site contenant cet ancêtre.

Pour finir, discutons rapidement les questions de modélisation abordées un peu plus haut. Puisque le taux de coalescence de plusieurs lignages assez proches est strictement positif, on a bien réussi à garder une population locale effective finie. Le caractère local des décès et reproductions est la base de la définition du processus ρ, et la dépendance en r de la fraction d'individus remplacée durant un événement permet d'adapter l'impact de cet événement en fonction de l'étendue de la zone touchée. Par exemple, des événements très locaux correspondront plutôt à la reproduction « normale » d'un individu et donc on pourra choisir une fraction d'individus remplacés plutôt faible ; en revanche, autoriser le rayon d'un événement a être très grand permet de modéliser des catastrophes plus rares pendant lesquelles on aura plutôt envie de remplacer une grande proportion de la population par les descendants d'un individu dont la famille recolonise très rapidement l'espace libéré.

2.3. Disparition de la structure dans la généalogie d'un modèle d'îles de Wright : résultats existants et contributions du chapitre 4.

2.3.1. *Phases de dispersion et de rassemblement de Wakeley.* Dans une série de travaux (voir [**Wak98, Wak99, WA01, Wak04**] en particulier), Wakeley considère un modèle d'îles de Wright dans lequel la population est constituée de D sous-populations et évolue en temps discret. L'île i contient un nombre N_i d'individus et reçoit M_i migrants à chaque génération, provenant de chacune des $D - 1$ autres îles avec même probabilité ; un mécanisme de type *Wright-Fisher* (chaque individu de la nouvelle génération indépendamment des autres choisit son parent uniformément au hasard parmi les $N_i + M_i$ parents possibles) ramène ensuite la population locale à N_i individus. Dans [**Wak04**], Wakeley inclut également dans ce scénario des événements d'extinction d'une fraction des communautés suivie par la recolonisation instantanée de chaque île éteinte (indépendamment les unes des autres), soit par les descendants d'un groupe de k individus provenant de la même île non-éteinte, soit par ceux de k individus choisis uniformément dans l'ensemble des communautés non-éteintes.

En faisant tendre le nombre d'îles vers l'infini, il obtient alors une décomposition de la généalogie de $n \ll D$ individus échantillonnés dans la population en deux phases :

- **une phase de dispersion**, pendant laquelle en un nombre $\mathcal{O}(1)$ de générations les lignages présents dans une même communauté fusionnent ou migrent vers une île différente, qui ne contient aucun des autres lignages que l'on suit avec une probabilité de l'ordre de $1 - \mathcal{O}(D^{-1})$. Cette phase se termine lorsque tous les lignages restants sont éparpillés sur des îles distinctes ;
- **une phase de rassemblement**, au cours de laquelle les lignages migrent d'une île à l'autre jusqu'à ce qu'au moins deux d'entre eux se croisent dans la même communauté et fusionnent ou se séparent à nouveau, puis reprennent leur déplacement. Par symétrie du mécanisme de migration, la probabilité qu'un lignage migrant atterrisse dans la même communauté que l'un des autres lignages est de l'ordre de $\mathcal{O}(D^{-1})$ et donc cette phase dure un nombre $\mathcal{O}(D)$ générations avant que le premier ancêtre commun à tout l'échantillon soit atteint.

Lorsque $D \to \infty$, la phase de dispersion (dont la longueur ne dépend que de la dynamique intra-île) devient donc beaucoup plus courte que celle de rassemblement, voire quasi-instantanée si on mesure le temps en unités de D générations. Sur cette échelle de temps, la généalogie de n individus traverse alors une phase très courte résultant en la coalescence de plusieurs lignages échantillonnés dans la même communauté et/ou leur dispersion sur des îles distinctes, puis évolue selon un coalescent de Kingman changé de temps par un paramètre effectif dépendant des différentes caractéristiques du modèle. Pour voir pourquoi cette seconde phase ne contient pas de coalescence multiple à la limite, calculons la probabilité que m lignages sur des îles distinctes fusionnent à la génération précédente. D'après le mécanisme d'évolution décrit dans le premier paragraphe, pour que ces lignages se retrouvent tous sur la même île il faut que chacun de leurs parents proviennent de la même communauté. Mais alors, soit ils descendent d'un migrant et donc ont choisi leur parent dans l'île en question avec probabilité $\mathcal{O}(D^{-1})$, soit ils ont été créés lors d'un événement d'extinction et puisque les recolonisateurs sont choisis indépendamment pour chaque île éteinte parmi la fraction d'îles non-affectées, on obtient à nouveau une probabilité de l'ordre de $\mathcal{O}(D^{-1})$ qu'ils proviennent d'une île en particulier. Au final, la probabilité que les $m-1$ derniers lignages se retrouvent sur l'île contenant le premier à la génération précédente est de l'ordre de $\mathcal{O}(D^{-(m-1)})$, qui est négligeable devant D^{-1} dès lors que $m \geq 3$: les coalescences multiples disparaissent bien à la limite.

L'outil principal utilisé par Wakeley pour formaliser cette convergence est le lemme suivant, qui apparaît régulièrement dans d'autres problèmes de séparation d'échelles de temps. Il est souvent attribué à Möhle, qui l'utilise dans [**Möh98**] pour étudier la généalogie de populations avec auto-fertilisation. Pour toute matrice $C = (c_{ij})$, notons $\|C\| = \max_i \sum_j |c_{ij}|$.

LEMME 1.39. *Soit* $(X_D)_{D\in\mathbb{N}}$ *une suite de chaînes de Markov à valeurs dans un même espace* S *et notons* Π_D *la matrice de transition de* X_D*. Supposons que*

1. $A := \lim_{D\to\infty} \Pi_D$ *existe et* $\Pi_D \neq A$ *pour* D *assez grand ;*
2. $P := \lim_{m\to\infty} A^m$ *existe ;*
3. $G := \lim_{D\to\infty} PB_DP$ *existe, où* $B_D := r_D(\Pi_D - A)$ *et* $r_D := \|\Pi_D - A\|^{-1}$.

Alors, si la suite $(X_D(0))_{D\in\mathbb{N}}$ *des valeurs initiales de* X_D *converge en loi vers une variable aléatoire de loi* μ*, les marginales de dimension finie du processus* $\{X_D(\lfloor r_D t\rfloor),\ t \geq 0\}$ *convergent* *en loi vers celles d'un processus X dont la valeur initiale $X(0)$ est distribuée selon μ et la matrice de transition au temps $t > 0$ est donnée par $\Pi(t) = P - I + e^{tG} = Pe^{tG} = e^{tG}P$.*

En essence, ce lemme nous dit que si on considère l'échelle de temps $r_D \gg 1$ sur laquelle B_D agit, l'évolution quasi-instantanée gouvernée par A peut être résumée en une projection sur la mesure stationnaire de A (P étant le projecteur associé) sur laquelle agit ensuite le générateur infinitésimal G correspondant à l'évolution plus lente. Notons qu'il n'est a priori pas possible de montrer une convergence plus forte que celle des marginales de dimension finie, car l'accumulation de transitions très rapides due à A empêche la suite de processus d'être tendue.

Wakeley utilise ce résultat en considérant les trois états dans lesquels deux lignages peuvent se trouver (sur la même île mais distincts, sur des îles différentes ou fusionnés) et en écrivant la matrice de transition entre ces trois états sous la forme $A + D^{-1}B + \mathcal{O}(D^{-2})$: A correspond alors à la dynamique intra-île rapide, B aux probabilités de rassemblement des deux lignages dû à leur migration vers la même île et le terme restant est un terme d'erreur disparaissant à la limite. Pour un nombre de lignages plus grand, l'ensemble des états possibles (et donc la taille de la matrice de transition) augmente très rapidement, si bien qu'il est plus raisonnable d'étudier simplement les phases de dispersion et de rassemblement séparément et de montrer que la phase de rassemblement ne comprend aucune collision multiple mais des fusions binaires à un taux indépendant de l'indice des îles distinctes sur lesquelles se trouvent les deux lignages impliqués.

2.3.2. *Résultats du chapitre 4 : critères de convergence vers un Ξ-coalescent dans le modèle d'îles de Wright.* Si l'approche décrite dans le paragraphe précédent est suffisante lorsque la probabilité que plus de deux lignages se rencontrent au même moment dans la même communauté tend vers 0 lorsque le nombre d'îles tend vers l'infini, elle ne l'est plus lorsque le mécanisme de recolonisation après un événement d'extinction massive autorise un très petit nombre d'individus à produire la fraction de la population éteinte à remplacer. Dans ce cas, la probabilité que plus de deux lignages aient leur parent dans la même communauté durant un tel événement devient d'ordre $\mathcal{O}(1)$ et une nouvelle phase très rapide de dispersion survient avant que les lignages soient à nouveau éparpillés sur des îles différentes. Dans le chapitre 4, on opte pour des processus à temps continu, plus faciles à étudier grâce à leur caractérisation par un générateur infinitésimal donnant les taux auxquels les différents événements arrivent. On cherche alors des conditions sur le mécanisme de reproduction intra-île, de migration et d'extinction et recolonisation pour que, lorsque $D \to \infty$, la généalogie d'un nombre fini d'individus associée au modèle d'îles que l'on considère se décompose aussi en une alternance de phases instantanées de dispersion et de phases plus lentes de rassemblement et que le coalescent structuré qui la décrit converge sur l'échelle de temps la plus longue vers un coalescent à collisions multiples et simultanées non-structuré.

Plus précisément, supposons que l'évolution de la population dans le sens *normal* du temps suive le même mécanisme sur chaque île et que chaque communauté joue un rôle symétrique dans les migrations et les recolonisations. Ainsi, si on échantillonne n individus et que l'on retrace leur généalogie, seule la manière dont les lignages sont regroupés dans différentes (au plus n) communautés est importante et on peut donc travailler dans l'espace P_n^s des partitions structurées non-ordonnées défini de la manière suivante :

DÉFINITION 1.40. *Notons $\tilde{\mathrm{P}}_n^s$ l'ensemble*

$$\tilde{\mathrm{P}}_n^s := \big\{ \big(\{B_1, \ldots, B_{i_1}\}, \ldots, \{B_{i_{n-1}+1}, \ldots, B_{i_n}\}\big) \ : \ 0 \leq i_1 \leq \ldots \leq i_n \leq n, \\ \emptyset \neq B_j \subset [n] \ \forall j \in \{1, \ldots, i_n\}, \ \{B_1, \ldots, B_{i_n}\} \in \mathcal{P}_n \big\}$$

des n-uplets d'ensembles (dont certaines composantes peuvent être vides), et définissons une relation d'équivalence $\sim$ sur $\tilde{\mathrm{P}}_n^s$ par $\xi \sim \xi'$ ssi il existe une permutation σ de $[n]$ telle que si

$\mathcal{B}_1 := \{B_1, \dots, B_{i_1}\}$, $\dots$, $\mathcal{B}_n := \{B_{i_{n-1}+1}, \dots, B_{i_n}\}$ *sont les composantes du vecteur* ξ*, alors* $\xi' = (\mathcal{B}_{\sigma(1)}, \dots, \mathcal{B}_{\sigma(n)})$*. Le quotient de* $\tilde{\mathrm{P}}^s_n$ *par* $\sim$ *est noté* P^s_n*, et on appelle tout élément de* P^s_n *une* partition structurée non-ordonnée de $[n]$.

Pour nos applications, chaque composante $\mathcal{B}_j$ représente une île particulière contenant certains des lignages ancestraux de l'échantillon et les blocs B_k spécifient la partition de cet échantillon en fonction des relations généalogiques entre les individus au temps considéré. Les composantes vides permettent de travailler avec des vecteurs de taille constante égale à n, indépendamment du nombre D de communautés que l'on fera tendre vers l'infini par la suite.

Un processus markovien à valeurs dans P^s_n pour lequel les blocs de la partition de $[n]$ ne peuvent que fusionner ou changer de composante (*migrer*) est appelé *processus généalogique structuré*. Une réalisation possible d'un tel processus sur P^s_4 est donnée par :

$$\begin{aligned}\big(\{\{1\}\},\{\{2\}\},\{\{3\}\},\{\{4\}\}\big) &\to \big(\{\{1\},\{2\}\},\{\{3,4\}\},\emptyset,\emptyset\big)\\ &\to \big(\{\{1\}\},\{\{2\}\},\{\{3,4\}\},\emptyset\big)\\ &\to \big(\{\{1,2,3,4\}\},\emptyset,\emptyset,\emptyset\big).\end{aligned}$$

Dans cet exemple, chaque lignage est initialement seul dans son île. Durant la première transition, l'ancêtre de 1 ou de 2 migre et tous deux se retrouvent dans la même communauté mais ne fusionnent pas, tandis que la migration qui amène les ancêtres de 3 et 4 sur la même île est immédiatement suivie par la coalescence des deux lignées ancestrales. Un bloc ne pouvant pas se diviser, 3 et 4 restent ensuite dans le même bloc à tout temps. Lors de la deuxième transition, les lignages de 1 et 2 n'ayant pas fusionné se séparent géographiquement par la migration de l'un d'entre eux. Pour finir, les trois blocs constituant la partition ancestrale se retrouvent dans la même composante et fusionnent, atteignant ainsi la valeur (absorbante) finale du processus.

Fixons un nombre $n \in \mathbb{N}$ d'individus et considérons une suite $(\mathcal{A}^D)_{D\in\mathbb{N}}$ de processus généalogiques structurés à valeurs dans P^s_n, sujets aux événements suivants :

- des coalescences intra-îles et des migrations de lignages vers des îles *vides* (i.e., ne contenant pas d'autres lignages de l'échantillon) à un taux d'ordre $\mathcal{O}(r_D)$, où $r_D \to \infty$ lorsque $D \to \infty$;
- des regroupements de lignages dispersés, éventuellement suivis par la fusion de certains d'entre eux, à un taux d'ordre $\mathcal{O}(1)$.

Le générateur de $\mathcal{A}^D$ s'écrit donc sous la forme

$$G^D = r_D\Psi + \Gamma + R_D,$$

où Ψ, Γ et R_D sont des opérateurs linéaires bornés et $R_D \to 0$ lorsque $D \to \infty$. Les conditions énoncées dans le paragraphe 3.1 du chapitre 4 formalisent cette description.

Par analogie avec les résultats de Wakeley (et les théorèmes de perturbation du chapitre 1.7 de [**EK86**] pour les processus à temps continus, qui ne s'appliquent pas ici pour la raison discutée à la fin du paragraphe 3.2 du chapitre 4), on commence donc par s'intéresser à la phase rapide correspondant à Ψ. Dans le paragraphe 2.2 du même chapitre, on introduit le processus généalogique structuré ξ correspondant, dans lequel seuls des lignages présents dans la même île peuvent fusionner ou migrer vers des îles vides. Deux lignages situés dans des communautés différentes ne sont jamais rassemblés sur la même île par ce processus, qui correspond donc bien à la première phase de dispersion décrite par Wakeley. On montre ensuite que ξ partant d'une valeur $\zeta \in \mathrm{P}^s_n$ atteint en un temps p.s. fini un état absorbant aléatoire $\underline{\zeta}$ à valeurs dans Π_n, le sous-ensemble de P^s_n des partitions structurées de $[n]$ dont chaque composante contient au plus un bloc. La valeur finale $\underline{\zeta}$ sert alors de point de départ à la seconde phase, gouvernée par Γ.

Puisque l'évolution très rapide due à Ψ conduit à une valeur finale dans laquelle chaque lignage appartient à une communauté différente et que la longueur de la phase de rassemblement suivante (prenant fin lorsqu'au moins deux lignages dispersés se retrouvent sur une même île) est quant à elle macroscopique, on s'attend à ce que $\mathcal{A}^D$ passe un temps négligeable hors de Π_n. On définit donc dans le paragraphe 2.3 un processus $\mathcal{A}$ à valeurs dans Π_n, évoluant de la manière suivante : lorsque $\mathcal{A}_t = \chi \in \Pi_n$, Γ donne le taux et la nature des *collisions géographiques* rassemblant un ou plusieurs groupes de lignages de χ sur des îles identiques. Si $\zeta \in \mathrm{P}_n^s$ est le résultat de la première collision, la nouvelle valeur de $\mathcal{A}$ juste après cet événement est choisie selon la loi de $\underline{\zeta}$. Le résultat principal de ce chapitre est le Théorème 4.4, dans lequel on montre que les marginales de dimension finie de $\mathcal{A}^D$ partant de $\zeta \in \mathrm{P}_n^s$ convergent en loi vers celles de $\mathcal{A}$ partant de $\underline{\zeta}$, sauf au temps $t = 0$. La question de la tension de la suite $(\mathcal{A}^D)_{D\in\mathbb{N}}$ est abordée dans le paragraphe 3.3.

La preuve du Théorème 4.4 utilise les temps d'arrêts alternés $(\sigma_i^D)_{i\in\mathbb{N}}$ et $(\tau_i^D)_{i\in\mathbb{N}}$ auxquels $\mathcal{A}^D$ retourne dans Π_n puis subit une collision géographique. Si $(\sigma_i)_{i\in\mathbb{N}}$ est la suite de temps auxquels une collision géographique arrive au processus limite $\mathcal{A}$, on commence par montrer que pour tout $i \in \mathbb{N}$, la loi du couple $(\sigma_i^D, \mathcal{A}^D_{\sigma_i^D})$ sur l'événement $\{\sigma_i^D < \infty\}$ converge vers celle de $(\sigma_i, \mathcal{A}_{\sigma_i})$ sur l'événement $\{\sigma_i < \infty\}$, et que le couple $(\tau_i^D, \mathcal{A}^D_{\tau_i^D})$ sur $\{\tau_i^D < \infty\}$ converge en loi vers $(\sigma_{i+1}, \chi(\mathcal{A}_{\sigma_i}))$ sur $\{\sigma_{i+1} < \infty\}$, où $\chi(\zeta)$ est la variable aléatoire donnant le résultat de la première collision géographique subie par $\mathcal{A}$ lorsque $\mathcal{A}_0 = \zeta \in \Pi_n$. En particulier, la longueur $\sigma_i^D - \tau_{i-1}^D$ de la i-ème période que $\mathcal{A}^D$ passe en dehors de Π_n converge en probabilité vers 0. On peut alors en déduire la convergence en loi des marginales unidimensionnelles de $\mathcal{A}^D$ puis par récurrence celle des marginales fini-dimensionnelles. On répond ainsi à la première question que l'on s'était posée, à savoir les conditions de convergence des processus généalogiques structurés vers un processus alternant des phases instantanées de dispersion et des phases de longueur $\mathcal{O}(1)$ de rassemblement.

La question suivante est celle de la disparition de la structure dans la généalogie limite. Le Lemme 4.12 et la Proposition 4.16 donnent des conditions nécessaires et suffisantes sur les taux de Ψ et Γ pour que la généalogie sur les échelles de temps respectivement rapide et lente soient cohérentes. Sous ces conditions, le processus $\mathcal{A}^u$ à valeurs dans $\mathcal{P}_n$ induit par $\mathcal{A}$ (obtenu en « oubliant » la structure géographique des lignages) est un Ξ-coalescent que l'on peut caractériser en fonction des différents taux des collisions géographiques et de la distribution du résultat de la phase de dispersion qui les suit. La Proposition 4.25 montre réciproquement que pour obtenir une généalogie cohérente sur l'échelle de temps rapide (celle générée par Ψ) et un coalescent à collisions multiples et simultanées non-spatial sur l'échelle de temps plus longue, il est nécessaire d'imposer les différentes conditions décrites dans le paragraphe 3.1.

2.4. Disparition de la structure dans la généalogie de modèles *stepping stone* et Λ-Fleming-Viot : résultats existants et contributions du chapitre 5.

2.4.1. *Généalogies d'individus initialement éloignés dans le modèle stepping stone.* Cette disparition de la structure dans la généalogie d'une population spatiale apparaît dans d'autres cadres que le modèle d'îles de Wright. Sous certaines hypothèses, le modèle stepping stone se prête également bien à ce phénomène, comme l'ont établi plusieurs travaux que nous passons en revue dans ce paragraphe.

La séparation d'échelle de temps qui nous intéresse à présent est celle qui se produit lorsqu'on échantillonne des individus dans des communautés très éloignées les unes des autres : le temps que deux lignages mettent alors à se rencontrer est tellement long que

- chaque paire de lignages a à peu près la même probabilité d'être la première à se retrouver dans la même communauté ;

– lorsqu'une telle rencontre arrive, le temps supplémentaire que mettent les deux lignages en question à fusionner est négligeable par rapport à leur temps de rencontre et peut être considéré comme instantané lorsque la distance initiale entre les lignages échantillonnés tend vers l'infini.

L'un des premiers travaux à avoir mis en avant cette décomposition est [**CG86**], où Cox et Griffeath considèrent le processus dual d'un modèle du votant sur $\mathbb{Z}^2$, qui dans leur cas n'est autre qu'un système de marches aléatoires aux plus proches voisins qui fusionnent dès qu'elles se trouvent sur le même site. Ils montrent alors que si n particules partent d'une distance $\mathcal{O}(L)$ les unes des autres, le processus η^L défini par

$$\eta_t^L := \# \{\text{particules distinctes au temps } L^{2t}\}, \qquad t \geq 1,$$

converge au sens de la convergence en loi des marginales de dimension finie lorsque $L \to \infty$ vers le processus $\{\eta_{\log t},\ t \geq 1\}$, où $\{\eta_s,\ s \geq 0\}$ compte le nombre de blocs d'un coalescent de Kingman à valeurs dans $\mathcal{P}_n$ et de valeur initiale la partition triviale de $[n]$. En particulier, la probabilité que deux marches aléatoires initialement à distance xL fusionnent avant le temps L^{2t} ($t \geq 1$) tend vers $1 - e^{-\log t} = 1 - t^{-1}$. Notons que cette convergence a lieu indépendamment de la distance initiale précise entre les particules, du moment qu'elle est de l'ordre de L. Cox et Griffeath complètent ce résultat dans [**CG90**] en montrant que sous les mêmes hypothèses, les marginales de dimension finie du processus « généalogique » associé (dont la valeur à un temps s est une partition regroupant dans chaque bloc les étiquettes des particules ayant fusionné avant le temps s) sur la même échelle de temps exponentielle L^{2t} convergent vers celles du coalescent de Kingman parcouru à vitesse $\log t$, lorsque $L \to \infty$.

Dans un autre article, Cox s'intéresse à nouveau au modèle du votant, mais défini cette fois sur un tore de côté L inclus dans $\mathbb{Z}^d$. En considérant le processus dual de marches aléatoires coalescentes partant de n particules initialement à distance $\mathcal{O}(L)$ les unes des autres, il montre dans le Théorème 5 de [**Cox89**] que les marginales unidimensionnelles (et par une extension facile celles de dimension finie) du processus η^L donné en dimension 2 par

$$\eta_t^L := \# \{\text{particules distinctes au temps } (L^2 \log L)t\}, \qquad t \geq 0,$$

convergent en loi lorsque $L \to \infty$ vers celles de $\{\eta_t, t \geq 0\}$ introduit au paragraphe précédent. La différence dans la manière de renormaliser le temps entre le cas où l'on travaille sur $\mathbb{Z}^2$ et celui où l'on se place sur un tore discret de côté L vient du fait que limiter l'espace disponible pour une particule permet à deux marches aléatoires indépendantes de se rencontrer en un temps beaucoup plus court, de l'ordre de $L^2 \log L$ en dimension 2 (ce temps est d'ordre L^2 si $d = 1$ et L^d si $d \geq 3$).

Toujours dans la même veine, Cox et Durrett dans [**CD02**] puis Zähle, Cox et Durrett dans [**ZCD05**] étudient la généalogie associée à un modèle stepping stone sur un tore de côté L inclus dans $\mathbb{Z}^2$, dans lequel chaque site contient un nombre fini d'individus se reproduisant à l'intérieur de leur communauté et envoyant des migrants vers d'autres îles. Comme on l'a vu dans le paragraphe 2.2.1, le processus dual est encore un système de marches aléatoires coalescentes ; les fusions de lignages présents sur une même île ne sont plus instantanées comme dans le cas du modèle du votant, mais arrivent à un taux inversement proportionnel à la taille de la population dans les îles. La marche aléatoire effectuée par un lignage n'est pas nécessairement aux plus proches voisins, mais on suppose que ses sauts sont symétriques et de variance σ^2 finie, indépendante de L. En employant des techniques similaires à celles de [**Cox89**], ils montrent entre autres que lorsque la taille de chaque communauté reste constante mais que L tend vers l'infini, la généalogie de n individus échantillonnés uniformément parmi les L^2 îles et considérée sur l'échelle de temps $L^2 \log L$ converge vers un coalescent de Kingman changé de temps par un facteur dépendant de

σ^2, mais pas de la taille des populations locales. L'énoncé précis concerne le processus comptant le nombre de lignages distincts au temps $(L^2 \log L)t$ dans le passé, mais les mêmes arguments que dans [**CG90**] donneraient sans plus d'effort la convergence de la suite de processus généalogiques structurés des n individus vers un coalescent de Kingman.

2.4.2. *Résultats du chapitre 5 : généalogies du processus Λ-Fleming-Viot spatial sur un très grand tore.* Remplaçons le modèle stepping stone du paragraphe précédent par un processus Λ-Fleming-Viot spatial ρ (cf. Définition 1.35) et le tore discret en dimension 2 par le tore $\mathbb{T}(L) \subset \mathbb{R}^2$ de côté L. Si on suppose que les rayons des événements sont bornés par une constante R indépendante de L, il n'est pas difficile de se convaincre que l'évolution de ρ vue à l'échelle « macroscopique » du tore entier ressemble beaucoup à celle d'un modèle stepping stone dans lequel la distance parcourue par un migrant est bornée. En particulier, deux lignages situés à une distance supérieure à $2R$ ne peuvent pas être pris dans le même événement du processus ponctuel de Poisson donnant l'évolution de ρ ; ils évoluent suivant la loi de processus de Lévy indépendants (des processus de Poisson composés symétriques dont la variance des sauts est bornée, pour être précis) jusqu'à ce qu'ils se retrouvent à une distance inférieure à $2R$. Les corrélations entre les mouvements de différents lignages et les coalescences sont donc très locales, tout comme pour le dual du modèle du votant. Par conséquent, on peut s'attendre au même type de convergence que dans [**ZCD05**], à savoir que la généalogie de n individus échantillonnés uniformément sur $\mathbb{T}(L)$ considérée sur l'échelle de temps $L^2 \log L$ converge lorsque $L \to \infty$ vers un coalescent de Kingman.

Pour avoir une chance de faire apparaître des coalescences multiples dans le processus généalogique limite, fixons $\alpha \in (0, 1]$ et supposons que le processus Λ-Fleming-Viot spatial sur $\mathbb{T}(L)$ évolue grâce à deux types d'événements :

- des *petits* événements, donnés par un processus ponctuel de Poisson Π_L^s sur $\mathbb{R} \times \mathbb{T}(L) \times (0, \infty)$ d'intensité $dt \otimes dx \otimes \mu^s(dr)$. Si $(t, x, r) \in \Pi_L^s$, le centre de l'événement est x, son rayon est r et la fraction d'individus remplacés à chaque site affecté est distribuée selon une mesure de probabilité ν_r^s.
- des *grands* événements, donnés par un processus ponctuel de Poisson Π_L^B sur $\mathbb{R} \times L^{-\alpha}\mathbb{T}(L) \times (0, \infty)$ d'intensité $\rho(L)^{-1} dt \otimes dx \otimes \mu^B(dr)$, où $(\rho(L))_{L \in \mathbb{N}}$ est une suite croissante de réels positifs tendant vers $+\infty$. Si $(t, x, r) \in \Pi_L^B$, le centre de l'événement est $L^\alpha x$, son rayon est $L^\alpha r$ et la fraction d'individus remplacés à chaque site affecté est distribuée selon une mesure de probabilité ν_r^B.

On autorise $\rho(L)$ à prendre des valeurs infinies, auquel cas les grands événements n'arrivent jamais. Supposons en outre que les supports de μ^s et μ^B sont bornés par deux constantes R^s et R^B respectivement. Un calcul rapide utilisant (1.31) montre que sous ces hypothèses, un lignage ancestral effectue des sauts de taille $\mathcal{O}(1)$ à taux $\mathcal{O}(1)$ et des sauts de taille $\mathcal{O}(L^\alpha)$ à taux $\mathcal{O}(\rho(L)^{-1})$.

L'objectif du chapitre 5 est de déterminer la généalogie limite de n individus échantillonnés uniformément sur $\mathbb{T}(L)$ (ou de manière équivalente, à distance $\mathcal{O}(L)$ les uns des autres) lorsque la taille L du tore tend vers l'infini. A cause des grands événements, l'échelle de temps à considérer n'est plus nécessairement $L^2 \log L$. Par ailleurs, on obtient des processus limites différents selon que $\alpha < 1$ ou $\alpha = 1$.

Le cas $\alpha < 1$

En utilisant une technique similaire à celle de [**CD02**], on montre dans la Proposition 5.26 que le temps nécessaire à deux lignages partant d'une distance $\mathcal{O}(L)$ pour se retrouver pour la première fois à une distance inférieure à $2R^B L^\alpha$ (avant quoi ils se déplacent de manière i.i.d.) est de l'ordre de

$$\varpi_L := \begin{cases} \frac{(1-\alpha)\rho(L)L^2 \log L}{2\pi\sigma_B^2 L^{2\alpha}} & \text{si } \rho(L) \ll L^{2\alpha}, \\ \frac{(1-\alpha)L^2 \log L}{2\pi(\sigma_s^2 + b\sigma_B^2)} & \text{si } \rho(L)^{-1} L^{2\alpha} \to b \in [0,\infty), \end{cases}$$

où σ_s^2 et σ_B^2 s'expriment en fonction de l'intensité des petits et grands sauts d'un lignage respectivement. Ensuite, dans le cas où $\rho(L) \ll L^{2\alpha} \log L$, les grands événements sont assez fréquents pour causer la coalescence des deux lignages ainsi rassemblés en un temps négligeable devant ϖ_L : la Proposition 5.28 assure que le temps que deux lignages initialement à distance au plus $2R^B L^\alpha$ mettent à fusionner est plus petit que $\varpi_L/(\log L)^2$ avec une probabilité tendant vers 1. Sinon, ce sont les petits événements qui causent la fusion des deux lignages, en un temps négligeable par rapport au temps d'ordre

$$\varpi'_L := \frac{L^2 \log L}{2\pi\sigma_s^2}$$

qu'ils mettent à se retrouver à une distance inférieure à $2R^s$ pour la première fois (avant quoi les grands événements sont assez rares pour que l'on puisse faire l'approximation que les lignages se déplacent de manière i.i.d. et n'ont aucune chance de fusionner).

On montre ensuite dans le Lemme 5.32 que lorsque deux lignages fusionnent, la probabilité que les autres lignages se trouvent à distance au moins $L/\log L$ les uns des autres et de la paire qui fusionne tend vers 1, et donc la généalogie limite ne contient que des coalescences binaires. Par symétrie, chaque paire de lignages a asymptotiquement la même probabilité d'être la première à se retrouver à une distance lui permettant de fusionner, ce qui au final nous donne le résultat suivant (cf. Théorème 5.11). Notons $\mathcal{A}^L$ le processus à valeurs dans les partitions marquées de $[n]$ qui retrace la généalogie de n individus échantillonnés à distance d'ordre $\mathcal{O}(L)$ les uns des autres sur $\mathbb{T}(L)$, et $\mathcal{A}^{L,u}$ sa restriction à $\mathcal{P}_n$ obtenue en oubliant la structure spatiale.

THÉORÈME 1.41. *Le processus (non-markovien) $\mathcal{A}^{L,u}$, considéré sur l'échelle de temps ϖ_L si $\rho(L) \ll L^{2\alpha} \log L$ et ϖ'_L sinon, converge en loi en tant que processus à trajectoires dans $D_{\mathcal{P}_n}[0,\infty)$ vers la restriction du coalescent de Kingman à $\mathcal{P}_n$.*

REMARQUE 1.42. *Si $\rho(L)$ n'est pas négligeable devant $L^{2\alpha} \log L$, on ne peut en fait montrer ce résultat que lorsque $L^{4\alpha}/\rho(L)$ est borné ou lorsque $\rho(L) \gg L^2 \log L$. En effet, si $\rho(L) \approx L^{2\alpha} \log L$, une fois que deux lignages sont rassemblés à distance $2R^B L^{2\alpha}$ les petits et les grands événements peuvent causer leur fusion en un temps comparable ; on s'attend alors à ce que la généalogie de l'échantillon dépende des trajectoires précises des lignages. Dans le cas restant, les mêmes arguments que précédemment montrent que les fusions ne peuvent être que binaires et causées par les petits événements, mais on est incapable d'obtenir le comportement asymptotique du temps de rencontre à distance $2R^s$ de deux lignages.*

Le cas $\alpha = 1$

Contrairement au cas précédent où les grands événements restaient de taille très petite par rapport à celle du tore, à présent les événements générés par Π_L^B couvrent une fraction non-négligeable de $\mathbb{T}(L)$. Les lignages ancestraux évoluent alors par petits sauts rapides dus aux événements de Π_L^s, entre les grands sauts et les fusions plus rares dus à Π_L^B. Le processus généalogique limite dépend donc du rapport entre $\rho(L)$ et L^2 : on montre dans le Théorème 5.15 que

1. Si $\lim_{L\to\infty} L^{-2}\rho(L) = b \geq 0$, le processus généalogique structuré $\mathcal{A}^L$ dont les marques des blocs sont renormalisées par L^{-1} converge sur l'échelle de temps $\rho(L)$ vers un système de mouvements browniens sur $\mathbb{T}(1)$, évoluant de manière indépendante et à vitesse b entre les grands événements, et fusionnant ou sautant uniquement grâce à ces derniers.
2. Si $L^2 \ll \rho(L) \leq CL^2 \log L$, le processus $\mathcal{A}^{L,u}$ considéré sur l'échelle de temps $\rho(L)$ converge vers la restriction à $\mathcal{P}_n$ d'un Λ-coalescent dont les collisions multiples sont causées par les grands événements et la partie Kingman, présente uniquement si $\rho(L)$ est d'ordre $\mathcal{O}(L^2 \log L)$, est due aux petits événements.
3. Si $\rho(L) \gg L^2 \log L$, le processus $\mathcal{A}^{L,u}$ considéré sur l'échelle de temps $L^2 \log L$ converge vers un coalescent de Kingman (les grands événements sont beaucoup trop rares et les petits événements réduisent la généalogie de n individus à leur ancêtre commun avant que le premier grand événement se produise).

Part 1

Convergence of spatial branching processes among Poissonian obstacles

CHAPTER 2

Super-Brownian motion among hard obstacles

We prove a convergence theorem for a sequence of super-Brownian motions moving among hard Poissonian obstacles, when the intensity of the obstacles grows to infinity but their diameters shrink to zero in an appropriate manner. The superprocesses are shown to converge in probability for the law $\mathbf{P}$ of the obstacles, and $\mathbf{P}$-almost surely for a subsequence, towards a superprocess with underlying spatial motion given by Brownian motion and (inhomogeneous) branching mechanism $\psi(u,x)$ of the form $\psi(u,x) = u^2 + \kappa(x)u$, where $\kappa(x)$ depends on the density of the obstacles. This work draws on similar questions for a single Brownian motion. In the course of the proof, we establish precise estimates for integrals of functions over the Wiener sausage, which are of independent interest. The results of this chapter have been published in [**Véb09**].

Contents

1. Introduction

1.1. Superprocesses in random media. The purpose of this chapter is to investigate the behaviour of super-Brownian motion among random obstacles, when the density of these obstacles grows to infinity but their diameter shrinks to zero in an appropriate manner. More precisely, let us fix $d \geq 2$ and a domain D of $\mathbb{R}^d$, and let $c : \mathbb{R}^d \to [0,\infty)$ be a bounded measurable function. For every $\varepsilon \in (0, \frac{1}{2})$, let us define an obstacle configuration by

$$\Gamma_\varepsilon = \bigcup_{x \in \mathcal{P}^\varepsilon} \overline{B}(x,\varepsilon),$$

where $\mathcal{P}^\varepsilon$ is a Poisson point process on $\mathbb{R}^d$ with intensity $\log(\varepsilon^{-1})c(x)dx$ if $d=2$ and $\varepsilon^{2-d}c(x)dx$ if $d \geq 3$, and $\overline{B}(x,\varepsilon)$ denotes the closed ball of radius ε centered at x. This Poisson point process is defined on a probability space $(\mathbf{\Omega}, \mathcal{F}, \mathbf{P})$. On a different probability space, let us also consider a superprocess $\{X_t^\varepsilon,\ t \in [0,\infty)\}$ with critical branching mechanism $\psi(u) = u^2$ and underlying spatial motion given by Brownian motion killed when entering $D^c \cup \Gamma_\varepsilon$. Thus, for each ε, the superprocess X^ε can be seen as evolving in a random medium given by Γ_ε. A realization of $\{\Gamma_\varepsilon,\ \varepsilon \in (0,1/2)\}$ will be called an *environment*.

We wish to understand the behaviour of X^ε when ε tends to zero. As in most works about random media, two points of view can be adopted : either we fix an environment (quenched approach), or we average over the possible realizations of $\bigcup_{\varepsilon>0} \Gamma_\varepsilon$ (annealed approach). Although the results of this chapter are set in the quenched framework, the main ingredients of their proofs are "annealed-type" calculations. Moreover, the latter approach is also useful in obtaining a better understanding of where the scaling comes from and of what the limiting process might be. To simplify the analysis, let us first assume that $D = \mathbb{R}^d$ and let us consider a single Brownian motion ξ, independent of the obstacles. Denote by P_x the probability measure under which ξ starts from x. Let us define the random time T_ε as the entrance time of ξ into the set Γ_ε, that is

$$T_\varepsilon := \inf\{t \geq 0 : \xi_t \in \Gamma_\varepsilon\}.$$

In addition, for all $0 \leq s \leq t$, we denote by $S_\varepsilon(s,t)$ the Wiener sausage of radius ε along the time interval $[s,t]$, defined as

$$S_\varepsilon(s,t) = \{y \in \mathbb{R}^d : \inf_{s \leq r \leq t} |\xi_r - y| \leq \varepsilon\} = \bigcup_{r \in [s,t]} \big(\xi_r + \overline{B}(0,\varepsilon)\big).$$

The probability that the Brownian motion ξ hits Γ_ε before time t is equal to the probability that the centre of one of the obstacles lies in $S_\varepsilon(0,t)$. These centres are given by the Poisson point process $\mathcal{P}^\varepsilon$ and so, by averaging over the random obstacles and using Fubini's theorem, we obtain

$$\mathbf{E}\big[\mathrm{P}_0[T_\varepsilon > t]\big] = \mathrm{E}_0\big[\mathbf{P}[\mathcal{P}^\varepsilon \cap S_\varepsilon(0,t) = \emptyset]\big] = \mathrm{E}_0\left[\exp - s_d(\varepsilon) \int_{S_\varepsilon(0,t)} c(x)dx\right], \tag{2.1}$$

where

$$s_d(\varepsilon) = \begin{cases} \log(\varepsilon^{-1}) & \text{if } d = 2, \\ \varepsilon^{2-d} & \text{if } d \geq 3. \end{cases}$$

In the case $c = \nu$, the integral in (2.1) is just ν times the volume $\lambda(S_\varepsilon(0,t))$ of the Wiener sausage, whose asymptotics have been well studied owing to their connections with physical problems (see e.g. the introduction of [**Spi64**], [**KL74**] or [**DV75**]). Note that the large-t asymptotics of $\lambda(S_\varepsilon(0,t))$ are essentially equivalent to its small-ε asymptotics thanks to the equality in law :

$$\lambda(S_1(0,t)) \overset{(d)}{=} t^{d/2} \lambda(S_{t^{-1/2}}(0,1)).$$

A classical result of Kesten, Spitzer and Whitman (cf. [**IM65**], p.253) states that, if $d \geq 3$,

$$\lim_{\varepsilon \to 0} s_d(\varepsilon) \lambda(S_\varepsilon(0,t)) = k_d t \qquad \text{a.s.}, \tag{2.2}$$

where $k_d = \frac{(d-2)\pi^{d/2}}{\Gamma(d/2)}$ ($k_3 = 2\pi$) is the Newtonian capacity of the unit ball. The Kesten-Spitzer-Whitman convergence result was in fact stated for the large-time asymptotics of $\lambda(S_\varepsilon(0,t))$, but a scaling argument gives the previous statement, at least in the sense of convergence in probability. The convergence in (2.2) also holds if $d = 2$ (see [**LG86**]), with $k_2 = \pi$.

It is not hard to deduce from the preceding result that, at least when the function c is continuous,

$$\lim_{\varepsilon \to 0} s_d(\varepsilon) \int_{S_\varepsilon(0,t)} c(y)dy = k_d \int_0^t c(\xi_s)ds \qquad a.s. \tag{2.3}$$

It then follows from (2.1) that

$$\lim_{\varepsilon \to 0} \mathbf{E}\big[\mathrm{P}_0[T_\varepsilon > t]\big] = \mathrm{E}_0\left[\exp - k_d \int_0^t c(\xi_s)ds\right].$$

This argument, which is due to Kac [**Kac74**], can be interpreted in the following way. When ε tends to zero, the obstacles become dense in $\mathbb{R}^d$ (at least if the function c is everywhere positive), and the Brownian motion ξ_t gets absorbed in the obstacles at rate $k_d c(\xi_t)$.

Back to our initial problem about killed superprocesses, the result for a single Brownian particle suggests that the sequence X^ε should converge to the superprocess X^* with branching mechanism $\psi(u,x) = u^2 + k_d c(x)u$ and underlying spatial motion given by Brownian motion. We shall establish in this work that the distribution of X^* is, indeed, the limit of the distribution of X^ε as ε tends to 0, in $\mathbf{P}$-probability. Here, the distribution of X^ε is a probability measure on the Skorokhod space $D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$ of all càdlàg paths with values in $\mathcal{M}_f(\mathbb{R}^d)$ (the space of all finite measures on $\mathbb{R}^d$) and the preceding limit is in the sense of weak convergence. A stronger statement can be made, but only for subsequences: if the sequence ε_n decreases to 0 fast enough,

$$X^{\varepsilon_n} \overset{(d)}{\to} X^* \qquad \text{as } n \to \infty,\ \mathbf{P}\text{−a.s.}$$

Here, $\overset{(d)}{\to}$ denotes convergence in distribution. Let us emphasize the meaning of this result : except for a set of zero $\mathbf{P}$-measure, if we fix an environment, then the sequence of superprocesses X^{ε_n} evolving among these fixed obstacles converges in law to X^*. Theorem 2.1 and Corollary 2.2 are stated in a more general setting, allowing the superprocesses to live only within a domain D of $\mathbb{R}^d$.

The question we address in this chapter was motivated by analogous works on Brownian motion. An extensive literature is already available on this topic, reviewed for example in [**Szn98**]. Owing to the well-known properties of Poisson point processes, they seem to be a natural way to encode traps and have been frequently exploited in investigations of the behaviour of Brownian motion moving among “hard” obstacles, where the particle is killed instantaneously when hitting an obstacle as described above, or among “soft” obstacles, within which the Brownian particle is killed at a certain rate. Our approach is close to ideas developed by Kac in [**Kac74**], whose probabilistic method differs from the analytic method used by Papanicolaou and Varadhan [**PV80**] in a similar context. Both derive the convergence in the $L^2(\mathbf{P})$-norm of the semigroup of Brownian motion among random obstacles when the number of obstacles tends to infinity but their diameters tend to 0 (recall that $\mathbf{P}$ denotes the probability measure on the space where the obstacles are defined). Subsequently, Brownian motion among traps was studied in different settings, in particular by Sznitman, who devised the powerful method of enlargement of obstacles (see [**Szn98**]).

The problem of super-Brownian motion or branching Brownian motion among random obstacles was addressed recently by Engländer in [**Eng00**], [**EdH03**] and [**Eng08**], the latter paper dealing with soft obstacles. However, Engländer considers the supercritical case (instead of critical super-Brownian motion as we do) and keeps the sizes of obstacles fixed. Within the obstacles, a particle does not die but branches at a slower rate. His interest is in the long-term asymptotics of the process and, in particular, the survival probability and the growth rate of the support. His techniques are mostly analytic, in contrast with the probabilistic tools of the present work.

1.2. Statement of the main result. Let us first introduce some notation and construct the sequence of superprocesses X^ε from the historical superprocess corresponding to a super-Brownian motion on $\mathbb{R}^d$, independent of the obstacles. We refer to [**DP91**] for more details on historical superprocesses and their applications. If E is a topological space, $\mathcal{M}_f(E)$ stands for the space of all finite Borel measures on E.

The (Brownian) historical superprocess can be defined as follows. Let $\mathcal{W}$ be the set of all finite continuous paths in $\mathbb{R}^d$, and note that $\mathbb{R}^d$ can be viewed as a subset of $\mathcal{W}$ by identifying x with the path of length zero and initial point x. Then, let $\tilde{\xi}$ be the continuous Markov process in $\mathcal{W}$ whose transition kernel is described as follows: If $\tilde{\xi}_0 = (w(r),\ 0 \le r \le s) \in \mathcal{W}$, the law of $\tilde{\xi}_t$ is the law under $\mathrm{P}_{w(s)}$ of the concatenation of the paths $(w(r),\ 0 \le r \le s)$ and $(\xi_r,\ 0 \le r \le t)$. The historical superprocess H is defined as the superprocess on $\mathcal{W}$ with branching mechanism $\psi(u) = u^2$ and underlying spatial motion given by $\tilde{\xi}$. Thus, H takes values in $\mathcal{M}_f(\mathcal{W})$. The super-Brownian motion X^0 on $\mathbb{R}^d$, starting at $\mu \in \mathcal{M}_f(\mathbb{R}^d)$, can then be recovered from the historical superprocess starting at μ (which is viewed as a finite measure on the paths of length zero) through the formula

$$\langle X_t^0, f\rangle = \int_{\mathcal{W}} H_t(dw) f(w(t))$$

for all f bounded and measurable and all $t \geq 0$. Here, $\langle \nu, f\rangle$ denotes the integral of f against the measure ν.

We exploit this correspondence between the historical superprocess and super-Brownian motion further to construct the sequence of killed superprocesses which is of interest in this work. Let E be an open subset of $\mathbb{R}^d$, and recall the definition of the obstacle configuration Γ_ε. For every $\varepsilon > 0$, the superprocess $\{X_t^{\varepsilon,E}, t \in [0,\infty)\}$ is defined from the historical superprocess H via the formula

$$\langle X_t^{\varepsilon,E}, f\rangle = \int_{\mathcal{W}} H_t(dw) f(w(t)) \mathbb{I}_{\{\forall s\in[0,t],\ w(s)\in E\cap\Gamma_\varepsilon^c\}},$$

for all f bounded and measurable, and all $t \geq 0$. It is straightforward to verify that $X^{\varepsilon,E}$ is itself a super-Brownian motion with critical branching mechanism $\psi(u) = u^2$ and underlying spatial motion given by Brownian motion killed when entering $E^c \cup \Gamma_\varepsilon$. Furthermore, $X_0^{\varepsilon,E}$ is the restriction of μ to $E \cap \Gamma_\varepsilon^c$.

Recall that we defined $k_2 = \pi$ and $k_d = \frac{d-2}{\Gamma(d/2)}\pi^{d/2}$ for $d \geq 3$. We also introduce another superprocess $X^{*,E}$, with branching mechanism $\psi(u,x) = u^2 + k_d c(x) u$ and underlying spatial motion given by Brownian motion killed when it exits E.

In practice, E will be either D or a bounded open subset of D. When there is no ambiguity, we shall suppress the dependence on E in the notation. We choose a sequence ε_n such that $\sum_n |\log \varepsilon_n|^{-1} < \infty$ if $d = 2$, and $\sum_n \varepsilon_n |\log \varepsilon_n| < \infty$ if $d \geq 3$. For instance, we may fix $\alpha > 1$ and set $\varepsilon_n = \exp(-n^\alpha)$ if $d = 2$ and $\varepsilon_n = n^{-\alpha}$ if $d \geq 3$.

We shall use the following notation.

- $\mathbb{P}_\mu$ is the (quenched) probability measure under which H starts at $\mu \in \mathcal{M}_f(\mathbb{R}^d) \subset \mathcal{M}_f(\mathcal{W})$. By the preceding correspondence, each superprocess $X^{\varepsilon,E}$ then starts under $\mathbb{P}_\mu$ from the restriction of μ to $E \cap \Gamma_\varepsilon^c$. It will be convenient to assume that $X^{*,E}$ is also defined under $\mathbb{P}_\mu$ and starts from the restriction of μ to E.
- To simplify notation, $X^{(n),E}$ will be a shorthand for the killed superprocess with parameter ε_n, and $\mathbb{P}_\mu^{(n),E}$ will be its law under $\mathbb{P}_\mu$. Likewise, $\mathbb{P}_\mu^{\varepsilon,E}$ (resp. $\mathbb{P}_\mu^{*,E}$) will be the law of $X^{\varepsilon,E}$ (resp. $X^{*,E}$) under $\mathbb{P}_\mu$.
- For all $t \geq 0$ and $x \in \mathbb{R}^d$, $\mathrm{P}_{t,x}$ will be a probability measure under which a Brownian motion ξ on $\mathbb{R}^d$, independent of the obstacles, starts from x at time t.
- $T^E := \inf\{t \geq 0 : \xi_t \in E^c\}$, $T_\varepsilon := \inf\{t \geq 0 : \xi_t \in \Gamma_\varepsilon\}$ and $T_{(n)} = T_{\varepsilon_n}$.

We can now state our main result.

THEOREM 2.1. *For every $\mu \in \mathcal{M}_f(D)$, $\mathbf{P}$-a.s.*

$$\mathbb{P}_\mu^{(n),D} \Rightarrow \mathbb{P}_\mu^{*,D} \text{ as } n \to \infty,$$

where the symbol $\Rightarrow$ refers to the weak convergence of probability measures.

As an immediate corollary, we also have :

COROLLARY 2.2. *For every $\mu \in \mathcal{M}_f(D)$, the sequence $\mathbb{P}_\mu^{\varepsilon,D}$ converges in $\mathbf{P}$-probability to $\mathbb{P}_\mu^{*,D}$ as ε tends to zero. In other words, for every $\delta > 0$, there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon \leq \varepsilon_0$,*

$$\mathbf{P}\left[d\big(\mathbb{P}_\mu^{\varepsilon,D}, \mathbb{P}_\mu^{*,D}\big) > \delta\right] < \delta,$$

where d *is the Prohorov metric on* $\mathcal{M}_1(D_{\mathcal{M}_f(D)}[0,\infty))$ *(here,* $\mathcal{M}_1(D_{\mathcal{M}_f(D)}[0,\infty))$ *is the space of all probability measures on* $D_{\mathcal{M}_f(D)}[0,\infty)$*).*

The rest of the chapter is devoted to the proofs of Theorem 2.1 and Corollary 2.2. In Section 2, we prove certain estimates for the rate of convergence in (2.3), which are of independent interest. These estimates are a key ingredient of the proof of Lemma 2.8 in Section 3. Then, we fix a bounded open subset B of D and prove the almost sure convergence of the distribution of $X^{(n),B}$ in two steps. First, we show in Section 3 that to each $k-$tuple $(t_1,\dots,t_k)$, there corresponds a set of $\mathbf{P}$-measure zero outside which $(X^{(n),B}_{t_1},\dots,X^{(n),B}_{t_k})_{n\geq 1}$ converges in law to $(X^{*,B}_{t_1},\dots,X^{*,B}_{t_k})$. Second, we prove in Section 4 that, with $\mathbf{P}$-probability 1, the sequence of superprocesses $X^{(n),B}$ is tight in $D_{\mathcal{M}_f(D)}[0,\infty)$. In Section 5, we complete the proof for a general domain D. Starting with a bounded subset of D is required for technical reasons, to ensure the finiteness of certain integrals which appear in the proof.

2. Some estimates for the Wiener sausage

Let us define the set $\mathcal{B}_1$ as the set of all bounded Borel measurable functions c on $\mathbb{R}^d$ such that $||c|| \leq 1$, where $||c||$ denotes the supremum norm of c. We have the following result (we write E_x for $E_{0,x}$ in the rest of the section):

PROPOSITION 2.3. *For every* $t \geq 0$*, there exists a constant* $C = C(t)$ *such that for every* $\varepsilon \in (0,\frac{1}{2}]$*, if* $d = 2$*,*

$$\sup_{c\in\mathcal{B}_1}\sup_{x\in\mathbb{R}^2} \mathrm{E}_x\left[\left(|\log\varepsilon|\int_{S_\varepsilon(0,t)} c(y)dy - \pi\int_0^t c(\xi_s)ds\right)^2\right] \leq \frac{C}{|\log\varepsilon|^2},$$

and if $d \geq 3$*,*

$$\sup_{c\in\mathcal{B}_1}\sup_{x\in\mathbb{R}^d} \mathrm{E}_x\left[\left(\varepsilon^{2-d}\int_{S_\varepsilon(0,t)} c(y)dy - k_d\int_0^t c(\xi_s)ds\right)^2\right] \leq C\varepsilon^2|\log\varepsilon|^2.$$

REMARK 2.4. *In the case* $c = 1$*, the bounds of Proposition 2.3 follow from the known results for the fluctuations of the volume of the Wiener sausage* [**LG88**]*. However, it does not seem easy to derive Proposition 2.3 from the special case* $c = 1$*. Note that the latter case suggests that the bound* $C\varepsilon^2|\log\varepsilon|^2$ *could be replaced by* $C\varepsilon^2|\log\varepsilon|$ *if* $d = 3$ *and by* $C\varepsilon^2$ *if* $d \geq 4$*. These refinements will not be needed in our applications.*

Proof of Proposition 2.3 for $d \geq 3$. To simplify notation, we prove the desired bound only for $t = 1$. A scaling argument then gives the result for any $t \geq 0$. Let us set

$$h(\varepsilon) = \sup_{c\in\mathcal{B}_1}\sup_{x\in\mathbb{R}^d} \mathrm{E}_x\left[\left(\varepsilon^{2-d}\int_{S_\varepsilon(0,1)} c(y)dy - k_d\int_0^1 c(\xi_s)ds\right)^2\right].$$

As a first step, let us notice that

$$\int_{S_\varepsilon(0,1)} c(y)dy = \int_{S_\varepsilon(0,1/2)} c(y)dy + \int_{S_\varepsilon(1/2,1)} c(y)dy - \int_{S_\varepsilon(0,1/2)\cap S_\varepsilon(1/2,1)} c(y)dy.$$

Also,

$$\varepsilon^{2-d}\int_{S_\varepsilon(0,1/2)} c(y)dy - k_d\int_0^{1/2} c(\xi_s)ds = \frac{\varepsilon^{2-d}}{2^{d/2}}\int_{\tilde{S}_{\varepsilon\sqrt{2}}(0,1)} c\Big(\frac{z}{\sqrt{2}}\Big)dz - \frac{k_d}{2}\int_0^1 c\Big(\frac{\tilde{\xi}_s}{\sqrt{2}}\Big)ds,$$

where $\tilde{\xi}_s = \sqrt{2}\,\xi_{s/2}$ for all $s \geq 0$ and $\tilde{S}_\varepsilon(a,b)$ is the Wiener sausage associated to $\tilde{\xi}$. Since the function $\tilde{c}(z) = c\big(\frac{z}{\sqrt{2}}\big)$ also belongs to $\mathcal{B}_1$, we obtain that

$$\mathrm{E}_x\left[\left(\varepsilon^{2-d}\int_{S_\varepsilon(0,1/2)} c(y)dy - k_d\int_0^{1/2} c(\xi_s)ds\right)^2\right] \leq \frac{1}{4}h(\varepsilon\sqrt{2}).$$

Likewise, using the Markov property at time $\frac{1}{2}$ and the preceding argument, we have

$$\mathrm{E}_x\left[\left(\varepsilon^{2-d}\int_{S_\varepsilon(1/2,1)} c(y)dy - k_d\int_{1/2}^{1} c(\xi_s)ds\right)^2\right] \leq \frac{1}{4}h(\varepsilon\sqrt{2}).$$

On the other hand, we have $\lambda\big(S_\varepsilon(0,1/2)\cap S_\varepsilon(1/2,1)\big) = \lambda\big(S'_\varepsilon(0,1/2)\cap S''_\varepsilon(0,1/2)\big)$, where $\xi'_t = \xi_{1/2-t}-\xi_{1/2}$ and $\xi''_t = \xi_{1/2+t}-\xi_{1/2}$ for every $t\in[0,1/2]$, and $S'_\varepsilon(0,1/2)$, resp. $S''_\varepsilon(0,1/2)$, denotes the Wiener sausage with radius ε associated to ξ', resp. ξ'', along the time interval $[0,1/2]$. Since ξ' and ξ'' are independent Brownian motions, we can use the following consequence of Corollary 3-2 in [**LG86**], and of [**LG88**], p.1012: There exists a constant $K_1(d)>0$ such that for every $\varepsilon\in(0,1/2]$

$$\mathrm{E}\Big[\lambda\big(S_\varepsilon(0,1/2)\cap S_\varepsilon(1/2,1)\big)^2\Big] \leq \begin{cases} K_1\varepsilon^4, & d=3\\ K_1\varepsilon^8|\log\varepsilon|^2, & d=4\\ K_1\varepsilon^{2d}, & d\geq 5.\end{cases}$$

Coming back to the definition of $h(\varepsilon)$, and using the triangle inequality in L^2, the fact that

$$\mathrm{E}_x\left[\left(\int_{S_\varepsilon(0,1/2)\cap S_\varepsilon(1/2,1)} c(y)dy\right)^2\right] \leq \mathrm{E}\Big[\lambda\big(S_\varepsilon(0,1/2)\cap S_\varepsilon(1/2,1)\big)^2\Big]$$

and the preceding inequalities, we obtain

$$\begin{aligned} h(\varepsilon) &\leq \sup_{c\in\mathcal{B}_1}\sup_{x\in\mathbb{R}^d}\left\{\mathrm{E}_x\left[\left(\varepsilon^{2-d}\int_{S_\varepsilon(0,1/2)} c(y)dy + \varepsilon^{2-d}\int_{S_\varepsilon(1/2,1)} c(y)dy - k_d\int_0^1 c(\xi_s)ds\right)^2\right]^{1/2}\right.\\ &\qquad \left. +\mathrm{E}_x\left[\varepsilon^{4-2d}\left(\int_{S_\varepsilon(0,1/2)\cap S_\varepsilon(1/2,1)} c(y)dy\right)^2\right]^{1/2}\right\}^2\\ &\leq \left\{\left(\frac{1}{2}h(\varepsilon\sqrt{2}) + 2u(\varepsilon)\right)^{1/2} + K_1'\psi_d(\varepsilon)\right\}^2, \end{aligned} \tag{2.4}$$

where $\psi_d(\varepsilon)=\varepsilon$ (resp. $\varepsilon^2|\log\varepsilon|$, resp. ε^2) if $d=3$ (resp. $d=4$, resp. $d\geq 5$) and

$$\begin{aligned} u(\varepsilon) = \sup_{c\in\mathcal{B}_1}\sup_{x\in\mathbb{R}^d}\Bigg|\mathrm{E}_x\Bigg[&\left(\varepsilon^{2-d}\int_{S_\varepsilon(0,1/2)} c(y)dy - k_d\int_0^{1/2} c(\xi_s)ds\right)\\ &\times\left(\varepsilon^{2-d}\int_{S_\varepsilon(1/2,1)} c(y)dy - k_d\int_{1/2}^1 c(\xi_s)ds\right)\Bigg]\Bigg|. \end{aligned}$$

Applying the Markov property at time $\frac{1}{2}$, we have

$$u(\varepsilon) = \sup_{c\in\mathcal{B}_1}\sup_{x\in\mathbb{R}^d}\left|\mathrm{E}_x\left[\left(\varepsilon^{2-d}\int_{S_\varepsilon(0,1/2)} c(y)dy - k_d\int_0^{1/2} c(\xi_s)ds\right)v(\varepsilon,\xi_{1/2})\right]\right|,$$

where

$$v(\varepsilon,z) = \mathrm{E}_z\left[\varepsilon^{2-d}\int_{S_\varepsilon(0,1/2)} c(y)dy - k_d\int_0^{1/2} c(\xi_s)ds\right].$$

We now use the following lemma.

LEMMA 2.5. *There exists a constant $K_2 > 0$ such that for all $z \in \mathbb{R}^d$, $\varepsilon \in (0, \frac{1}{2}]$ and $c \in \mathcal{B}_1$*

$$|v(\varepsilon, z)| \le K_2\, \varepsilon.$$

We postpone the proof of Lemma 2.5 and complete the case $d \ge 3$ of the Proposition. By Lemma 2.5, we have

$$|u(\varepsilon)| \le K_2\, \varepsilon \sup_{c\in\mathcal{B}_1} \sup_{x\in\mathbb{R}^d} \mathrm{E}_x\Big[\Big(\varepsilon^{2-d}\int_{S_\varepsilon(0,1/2)} c(y)dy - k_d\int_0^{1/2} c(\xi_s)ds\Big)^2\Big]^{1/2} \le \frac{K_2}{2}\,\varepsilon h(\varepsilon\sqrt{2})^{1/2}.$$

From (2.4), we obtain for every $\varepsilon \in (0, \frac{1}{2}]$

$$h(\varepsilon) \le \Big(\Big(\frac{1}{2}h(\varepsilon\sqrt{2}) + K_2\varepsilon h(\varepsilon\sqrt{2})^{1/2}\Big)^{1/2} + K_1'\psi_d(\varepsilon)\Big)^2.$$

Let us set $g(\varepsilon) = \varepsilon^{-1}h(\varepsilon)^{1/2}$. We thus have for $\varepsilon \in (0, \frac{1}{2}]$:

$$g(\varepsilon) \le \big(g(\varepsilon\sqrt{2})^2 + \sqrt{2}K_2 g(\varepsilon\sqrt{2})\big)^{1/2} + K_1'\varepsilon^{-1}\psi_d(\varepsilon). \tag{2.5}$$

Fix $r \in (1/4, 1/2]$ and set $u_n = g(r2^{-n/2})$ for every integer $n \ge 0$. Rewriting (2.5) in terms of u_n and noting that $\varepsilon^{-1}\psi_d(\varepsilon) = 1$ if $d = 3$ and $\varepsilon^{-1}\psi_d(\varepsilon) = o(1)$ as $\varepsilon \to 0$ if $d \ge 4$, we obtain for a constant $K_1'' > 0$ (independent of n)

$$u_{n+1} \le (u_n^2 + \sqrt{2}K_2u_n)^{1/2} + K_1'' = u_n\Big(1 + \frac{\sqrt{2}K_2}{u_n}\Big)^{1/2} + K_1'' \le u_n + \frac{\sqrt{2}K_2}{2} + K_1''.$$

It follows that $u_n \le u_0 + n\big(K_2 2^{-1/2} + K_1''\big)$ for every $n \ge 0$, from which we can conclude that there exists a constant K_3 such that for all $\varepsilon \in (0, 1/2]$,

$$g(\varepsilon) \le K_3|\log\varepsilon|$$

and thus

$$h(\varepsilon) \le K_3^2\varepsilon^2|\log\varepsilon|^2.$$

□

Proof of Lemma 2.5. We may assume that $z = 0$, and we fix $c \in \mathcal{B}_1$ (the constant K_2 will not depend on c). First, we have

$$\mathrm{E}_0\Big[k_d\int_0^{1/2} c(\xi_s)ds\Big] = k_d\int_{\mathbb{R}^d} dy\, c(y)\int_0^{1/2}\frac{ds}{(2\pi s)^{d/2}}\exp\Big(-\frac{|y|^2}{2s}\Big). \tag{2.6}$$

Let us define the random times $\tau_\varepsilon(y)$ and $L_\varepsilon(y)$ for all $\varepsilon > 0$ and $y \in \mathbb{R}^d$ by

$$\tau_\varepsilon(y) = \inf\{t \ge 0 :\ |\xi_t - y| \le \varepsilon\},$$
$$L_\varepsilon(y) = \sup\{t \ge 0 :\ |\xi_t - y| \le \varepsilon\},$$

with the conventions that $\inf\emptyset = +\infty$ and $\sup\emptyset = 0$. We thus have

$$\mathrm{E}_0\Big[\int_{S_\varepsilon(0,\frac{1}{2})} c(y)dy\Big] = \int_{\mathbb{R}^d} dy\, c(y)\,\mathrm{P}_0\Big[\tau_\varepsilon(y) \le \frac{1}{2}\Big]$$

$$= \int_{\mathbb{R}^d} dy\, c(y)\,\mathrm{P}_0\Big[0 < L_\varepsilon(y) \le \frac{1}{2}\Big] + \int_{\mathbb{R}^d} dy\, c(y)\,\mathrm{P}_0\Big[\tau_\varepsilon(y) \le \frac{1}{2} < L_\varepsilon(y)\Big]. \tag{2.7}$$

On the one hand,

$$\Big|\int dy\, c(y)\,\mathrm{P}_0\Big[\tau_\varepsilon(y) \le \frac{1}{2} < L_\varepsilon(y)\Big]\Big| \le \int dy\,\mathrm{P}_0\Big[\tau_\varepsilon(y) \le \frac{1}{2} \le L_\varepsilon(y)\Big].$$

We have

$$\int dy\, \mathrm{P}_0\Big[\tau_\varepsilon(y) \leq \frac{1}{2} \leq L_\varepsilon(y)\Big] = \mathrm{E}_0\Big[\lambda\big(S_\varepsilon(0,1/2) \cap S_\varepsilon(1/2,\infty)\big)\Big] = \mathrm{E}_0\Big[\lambda\big(S_\varepsilon(0,1/2) \cap S'_\varepsilon(0,\infty)\big)\Big],$$

where S'_ε denotes the Wiener sausage associated to a Brownian motion ξ' independent of ξ and also started from 0 under P_0. If $d = 3$, it is easily checked that

$$\mathrm{E}_0\Big[\lambda\big(S_\varepsilon(0,1/2) \cap S'_\varepsilon(0,\infty)\big)\Big] = \mathcal{O}(\varepsilon^2) \tag{2.8}$$

(use the fact that $\mathrm{P}_0\big[y \in S'_\varepsilon(0,\infty)\big] = \frac{\varepsilon}{|y|} \wedge 1$, together with the bound (3.d) in [**LG88**]). If $d \geq 4$,

$$\mathrm{E}_0\Big[\lambda\big(S_\varepsilon(0,1/2) \cap S'_\varepsilon(0,\infty)\big)\Big] = \varepsilon^d \mathrm{E}_0\Big[\lambda\big(S_1(0,\varepsilon^{-2}/2) \cap S'_1(0,\infty)\big)\Big]$$

$$= \begin{cases} \mathcal{O}(\varepsilon^4 |\log\varepsilon|), & \text{if } d = 4, \\ \mathcal{O}(\varepsilon^d), & \text{if } d \geq 5 \end{cases} \tag{2.9}$$

by [**LG88**], p.1010.

Let $\nu_{\varepsilon,y}(dz)$ denote the equilibrium measure of the ball $\overline{B}(y,\varepsilon)$, that is the unique finite measure on the sphere $\partial B(y,\varepsilon)$ such that for every x with $|x-y| > \varepsilon$,

$$\mathrm{P}_x\big[\tau_\varepsilon(y) < \infty\big] = \int \nu_{\varepsilon,y}(dz) G(z-x),$$

where $G(z) = \int_0^\infty (2\pi s)^{-d/2} \exp(-|z|^2/2s) ds = c_d |y|^{2-d}$ is the Green function of d-dimensional Brownian motion (c_d is a constant depending only on d). By a classical formula of probabilistic potential theory (see [**PS78**], p.61-62) we have

$$\mathrm{P}_0\Big[0 < L_\varepsilon(y) \leq \frac{1}{2}\Big] = \int_0^{1/2} ds \int \nu_{\varepsilon,y}(dz) \frac{1}{(2\pi s)^{d/2}} \exp\Big(-\frac{|z|^2}{2s}\Big).$$

It is well known that $\nu_{\varepsilon,y} = k_d \varepsilon^{d-2} \pi_{\varepsilon,y}$, where $\pi_{\varepsilon,y}$ denotes the uniform distribution on the sphere of radius ε centered at y. Recalling (2.6), (2.7), (2.8) and (2.9), we can write

$$\begin{aligned} |v(\varepsilon,z)| &= \left| k_d \int_{\mathbb{R}^d} dy\, c(y) \int_0^{1/2} \frac{ds}{(2\pi s)^{d/2}} \Big\{ \int \pi_{\varepsilon,y}(dz) \big(e^{-|z|^2/2s} - e^{-|y|^2/2s}\big) \Big\} \right| + \mathcal{O}(\phi_d(\varepsilon)) \\ &\leq k_d \int_{\mathbb{R}^d} dy \int_0^{1/2} \frac{ds}{(2\pi s)^{d/2}} \Big\{ \int \pi_{\varepsilon,y}(dz) \Big| e^{-|z|^2/2s} - e^{-|y|^2/2s} \Big| \Big\} + \mathcal{O}(\phi_d(\varepsilon)) \end{aligned}$$

where $\phi_d(\varepsilon) = \varepsilon$ (resp. $\varepsilon^2 |\log\varepsilon|$, resp. ε^2) if $d = 3$ (resp. $d = 4$, resp. $d \geq 5$). It follows that

$$\begin{aligned} &\int_{|y| \leq 10\varepsilon} dy \int_0^{1/2} \frac{ds}{(2\pi s)^{d/2}} \int \pi_{\varepsilon,y}(dz) \Big| e^{-|z|^2/2s} - e^{-|y|^2/2s} \Big| \\ &\quad \leq 2 \int_{|z| \leq 11\varepsilon} dz \int_0^{1/2} \frac{ds}{(2\pi s)^{d/2}}\, e^{-|z|^2/2s} \leq 2 \int_{|z| \leq 11\varepsilon} dz\, G(z) = \mathcal{O}(\varepsilon^2). \end{aligned}$$

On the other hand, we can find constants C and C' such that if $|y| > 10\varepsilon$ and $|z-y| = \varepsilon$,

$$\left| \exp\Big(-\frac{|z|^2}{2s}\Big) - \exp\Big(-\frac{|y|^2}{2s}\Big) \right| \leq C \left| \frac{|z|^2 - |y|^2}{s} \right| \exp\Big(-\frac{|y|^2}{4s}\Big) \leq C' \varepsilon \frac{|y|}{s} \exp\Big(-\frac{|y|^2}{4s}\Big).$$

Thus, with a constant K which may vary from line to line, we have

$$\int_{|y| > 10\varepsilon} dy \int_0^{1/2} \frac{ds}{(2\pi s)^{d/2}} \int \pi_{\varepsilon,y}(dz) \left| \exp\Big(-\frac{|z|^2}{2s}\Big) - \exp\Big(-\frac{|y|^2}{2s}\Big) \right|$$

$$\begin{aligned}
&\leq K\varepsilon \int_{|y|>10\varepsilon} dy\, |y| \int_0^{1/2} ds\, s^{-d/2-1} \exp\Big(-\frac{|y|^2}{4s}\Big) \\
&= K\varepsilon \int_{|y|>10\varepsilon} dy\, |y|^{1-d} \int_0^{1/(2|y|^2)} ds'\, s'^{-d/2-1} e^{-1/4s'} \\
&\leq K\varepsilon.
\end{aligned}$$

Combining the above, the proof of Lemma 2.5 is complete. □

Proof of Proposition 2.3 for $d = 2$. Let us define

$$h(\varepsilon) = \sup_{c\in\mathcal{B}_1} \sup_{x\in\mathbb{R}^d} \mathrm{E}\Big[\Big(|\log\varepsilon| \int_{S_\varepsilon(0,1)} c(y)dy - \pi \int_0^1 c(\xi_s)ds\Big)^2\Big].$$

By Corollary 3-2 in [**LG86**], we have

$$\mathrm{E}\Big[\lambda\big(S_\varepsilon(0,\frac{1}{2}) \cap S_\varepsilon(\frac{1}{2},1)\big)^2\Big] \leq \frac{K_1}{|\log\varepsilon|^4}.$$

The same technique as in the previous case yields

$$h(\varepsilon) \leq \Big\{\Big(\frac{1}{2}h(\varepsilon\sqrt{2}) + h(\varepsilon\sqrt{2})^{1/2} \sup_{c\in\mathcal{B}_1} \sup_{x\in\mathbb{R}^2} |v(\varepsilon,z)|\Big)^{1/2} + \frac{\sqrt{K_1}}{|\log\varepsilon|}\Big\}^2, \tag{2.10}$$

where

$$v(\varepsilon,z) = \mathrm{E}_z\Big[|\log\varepsilon| \int_{S_\varepsilon(0,1/2)} c(y)dy - \pi \int_0^{1/2} c(\xi_s)ds\Big].$$

We now use the following result, whose rather technical proof is deferred to the Appendix : There exists a constant K_2 such that, for $\varepsilon \in (0,1/2]$,

$$\sup_{c\in\mathcal{B}_1} \sup_{z\in\mathbb{R}^d} |v(\varepsilon,z)| \leq \frac{K_2}{|\log\varepsilon|}. \tag{2.11}$$

Hence, if $g(\varepsilon) = |\log\varepsilon| h(\varepsilon)^{1/2}$, we have for $\varepsilon \in (0,1/2]$,

$$g(\varepsilon) \leq \Big(\frac{1}{2}\frac{(\log\varepsilon)^2}{(\log\varepsilon\sqrt{2})^2}\, g(\varepsilon\sqrt{2})^2 + K_2 \frac{|\log\varepsilon|}{|\log\varepsilon\sqrt{2}|}\, g(\varepsilon\sqrt{2})\Big)^{1/2} + \sqrt{K_1}. \tag{2.12}$$

From (2.12), we can use arguments similar to the case $d \geq 3$ to infer that the function $g(\varepsilon)$ is bounded over $(0,1/2]$. Thus, there exists a constant K_3 such that for all $\varepsilon \in (0,1/2]$,

$$h(\varepsilon) \leq \frac{K_3}{|\log\varepsilon|^2}.$$

□

3. Almost sure convergence of the finite-dimensional distributions of $X^{(n),B}$

In the following, we fix a bounded open subset B of D and consider only the superprocesses killed outside B. We therefore suppress the dependence on B in the notation. In particular, $T = T^B$.

The following proposition is the first step in the proof of Theorem 2.1.

PROPOSITION 2.6. ***(Convergence of the finite-dimensional distributions)*** *Let* $\mu \in \mathcal{M}_f(B)$, $p \in \mathbb{N}$ *and* $t_1 < \ldots < t_p \in [0,\infty)$. *Then, under* $\mathbb{P}_\mu$

$$(X^{(n)}_{t_1}, \ldots, X^{(n)}_{t_p}) \overset{(d)}{\to} (X^*_{t_1}, \ldots, X^*_{t_p})$$

as $n \to \infty$, *on a set of* $\mathbf{P}$*-probability* 1.

Proof of Proposition 2.6. We fix an environment. Let $p \in \mathbb{N}$, $0 \leq t_1 < \ldots < t_p$ and $f_1, \ldots, f_p \in \mathcal{B}_{b+}(\mathbb{R}^d)$ be measurable, nonnegative and bounded functions. In the following, we shall denote $(t_1, \ldots, t_p)$ by $\mathbf{t}$ and $(f_1, \ldots, f_p)$ by $\mathbf{f}$.

Let $\mu \in \mathcal{M}_f(B)$. Following the notation in [**Dy91b**], we have :

$$\mathbb{E}_\mu \left[\exp - \sum_{i=1}^p \langle X^\varepsilon_{t_i}, f_i \rangle \right] = \exp - \langle \mu, w^\varepsilon_0 \rangle,$$

$$\mathbb{E}_\mu \left[\exp - \sum_{i=1}^p \langle X^*_{t_i}, f_i \rangle \right] = \exp - \langle \mu, w^*_0 \rangle,$$

where $w^\varepsilon = (w^\varepsilon_t(x);\ t \geq 0,\ x \in B)$ and $w^* = (w^*_t(x);\ t \geq 0,\ x \in B)$ are the unique nonnegative solutions to the following integral equations: for all $x \in B$ and $t \geq 0$,

$$w^\varepsilon_t(x) + \mathrm{E}_{t,x}\left[\int_t^\infty ds\ w^\varepsilon_s(\xi_s)^2 \mathbb{I}_{\{s < T \wedge T_\varepsilon\}} \right] = \sum_{i=1}^p \mathrm{E}_{t,x}\Big[f_i(\xi_{t_i}) \mathbb{I}_{\{t_i < T \wedge T_\varepsilon\}} \Big], \tag{2.13}$$

$$w^*_t(x) + \mathrm{E}_{t,x}\left[\int_t^\infty ds\ \big(w^*_s(\xi_s)^2 + k_d c(\xi_s) w^*_s(\xi_s) \big)\ \mathbb{I}_{\{s<T\}} \right] = \sum_{i=1}^p \mathrm{E}_{t,x}\Big[f_i(\xi_{t_i})\ \mathbb{I}_{\{t_i < T\}} \Big], \tag{2.14}$$

where by convention $\mathrm{E}_{t,x}[f(\xi_s)] = 0$ if $s < t$. By the standard argument of the proof of the Feynman-Kac formula, the integral equation (2.14) for w^* is equivalent to

$$w^*_t(x) + \mathrm{E}_{t,x}\left[\int_t^\infty ds\ w^*_s(\xi_s)^2 e^{-k_d \int_t^s c(\xi_u) du}\ \mathbb{I}_{\{s<T\}} \right] = \sum_{i=1}^p \mathrm{E}_{t,x}\Big[f_i(\xi_{t_i}) e^{-k_d \int_t^{t_i} c(\xi_u) du}\ \mathbb{I}_{\{t_i<T\}} \Big]. \tag{2.15}$$

The equivalence of the two integral equations (2.14) and (2.15) corresponds to the well-known fact that super-Brownian motion with branching mechanism $\psi(u,x) = u^2 + k_d c(x) u$ can also be constructed as the superprocess with branching mechanism $\psi(u) = u^2$ and underlying spatial motion given by Brownian motion killed at rate $k_d c(x)$.

REMARK 2.7. *Since w^ε and w^* are nonnegative, (2.13) and (2.14) imply that w^*_t and w^ε_t are equal to zero whenever $t > t_p$ (recall that by convention, the right-hand side of (2.13) or (2.14) is zero when $t > t_p$). Likewise, $w^\varepsilon_t(x) = 0$ if $x \in \Gamma_\varepsilon \cap B$, for every $t \geq 0$.*

By integrating over B the difference between (2.13) and (2.15), we obtain :

$$\begin{aligned}
\int_B dx\ |w^\varepsilon_t(x) - w^*_t(x)| \leq & \int_B dx \left| \sum_{i=1}^p \mathrm{E}_{t,x}\Big[f_i(\xi_{t_i}) \mathbb{I}_{\{t_i<T\}} (\mathbb{I}_{\{t_i < T_\varepsilon\}} - e^{-k_d \int_t^{t_i} c(\xi_u) du}) \Big] \right| \\
& + \int_B dx \left| \mathrm{E}_{t,x}\left[\int_t^\infty ds\ \mathbb{I}_{\{s<T\}} (\mathbb{I}_{\{s<T_\varepsilon\}} - e^{-k_d \int_t^s c(\xi_u) du}) w^\varepsilon_s(\xi_s)^2 \right] \right| \\
& + \int_B dx \left| \mathrm{E}_{t,x}\left[\int_t^\infty ds\ e^{-k_d \int_t^s c(\xi_u) du} \mathbb{I}_{\{s<T\}} (w^*_s(\xi_s)^2 - w^\varepsilon_s(\xi_s)^2) \right] \right|.
\end{aligned} \tag{2.16}$$

Let us start with the third term in the right-hand side of (2.16). The functions w^ε and w^* are bounded by $C_\mathbf{f} := \sum_{i=1}^p \|f_i\|$, hence bounding $\mathbb{I}_{\{s<T\}}$ by $\mathbb{I}_B(\xi_s)$ and $e^{-k_d \int_t^s c(\xi_u) du}$ by 1 yields

$$\begin{aligned}
& \int_B dx \left| \mathrm{E}_{t,x}\left[\int_t^\infty ds\ e^{-k_d \int_t^s c(\xi_u) du} \mathbb{I}_{\{s<T\}} (w^*_s(\xi_s)^2 - w^\varepsilon_s(\xi_s)^2) \right] \right| \\
& \leq 2C_\mathbf{f} \int_B dx\ \mathrm{E}_{t,x}\left[\int_t^\infty ds\ \mathbb{I}_B(\xi_s) |w^*_s(\xi_s) - w^\varepsilon_s(\xi_s)| \right] \\
& = 2C_\mathbf{f} \int_t^\infty ds \int_{B \times B} dx\, dz\ |w^*_s(z) - w^\varepsilon_s(z)|\ p_{s-t}(x,z)
\end{aligned}$$

$$\leq 2C_{\mathbf{f}} \int_t^\infty ds \int_B dz\, |w_s^*(z) - w_s^\varepsilon(z)|. \tag{2.17}$$

In the preceding estimates, $p_r(\cdot,\cdot)$ denotes the transition density at time r of d-dimensional Brownian motion. The last inequality stems from the observation that

$$\int_B p_{s-t}(x,z)dx = \int_B p_{s-t}(z,x)dx \leq 1.$$

We next show that the first two terms of (2.16) converge towards 0 $\mathbf{P}$-a.s. The key ingredient is the following result:

LEMMA 2.8. *Let $t_1 \in [0,\infty)$ and let $f \in \mathcal{B}_{b+}(\mathbb{R}^d)$ be a bounded nonnegative measurable function. Then, there exists a constant $K = K(c,t_1,d)$ such that, for every $t \in [0,\infty)$, $x \in B$ and $\varepsilon \in (0,1/2)$, if $d = 2$*

$$\mathbf{E}\Big[\mathrm{E}_{t,x}\Big[f(\xi_{t_1})\mathbb{I}_{\{t_1<T\}}\big(\mathbb{I}_{\{t_1<T_\varepsilon\}} - e^{-\pi\int_t^{t_1} c(\xi_u)du}\big)\Big]^2\Big] \leq K\,\|f\|^2\,\frac{1}{|\log\,\varepsilon|},$$

and if $d \geq 3$,

$$\mathbf{E}\Big[\mathrm{E}_{t,x}\Big[f(\xi_{t_1})\mathbb{I}_{\{t_1<T\}}\big(\mathbb{I}_{\{t_1<T_\varepsilon\}} - e^{-k_d\int_t^{t_1} c(\xi_u)du}\big)\Big]^2\Big] \leq K\,\|f\|^2\,\varepsilon|\log\,\varepsilon|.$$

The proof of Lemma 2.8 is postponed until the end of the section. Let us temporarily fix $t \in [0,t_p]$. Applying the lemma with $\varepsilon = \varepsilon_n$, we obtain for every $\delta > 0$ and every $i \in \{1,\dots,p\}$

$$\begin{aligned}
&\mathbf{P}\Big[\int_B\Big|\mathrm{E}_{t,x}\Big[f_i(\xi_{t_i})\mathbb{I}_{\{t_i<T\}}(\mathbb{I}_{\{t_i<T_{(n)}\}} - e^{-k_d\int_t^{t_i} c(\xi_u)du})\Big]\Big|\,dx > \delta\Big]\\
&\leq \frac{1}{\delta^2}\,\mathbf{E}\Big[\Big(\int_B\Big|\mathrm{E}_{t,x}\Big[f_i(\xi_{t_i})\mathbb{I}_{\{t_i<T\}}(\mathbb{I}_{\{t_i<T_{(n)}\}} - e^{-k_d\int_t^{t_i} c(\xi_u)du})\Big]\Big|\,dx\Big)^2\Big]\\
&\leq \frac{\lambda(B)}{\delta^2}\int_B \mathbf{E}\Big[\mathrm{E}_{t,x}\Big[f_i(\xi_{t_i})\mathbb{I}_{\{t_i<T\}}(\mathbb{I}_{\{t_i<T_{(n)}\}} - e^{-k_d\int_t^{t_i} c(\xi_u)du})\Big]^2\Big]dx\\
&\leq \begin{cases} \lambda(B)^2K\,\|f_i\|^2\delta^{-2}\,|\log\varepsilon_n|^{-1} & \text{if } d=2,\\ \lambda(B)^2K\,\|f_i\|^2\delta^{-2}\,\varepsilon_n|\log\varepsilon_n| & \text{if } d\geq 3,\end{cases}
\end{aligned}$$

which is summable by our assumptions on $(\varepsilon_n)_{n\geq 1}$. Hence, by the Borel-Cantelli lemma,

$$\mathbf{P}\text{-a.s.},\ \int_B\Big|\mathrm{E}_{t,x}\Big[f_i(\xi_{t_i})\mathbb{I}_{\{t_i<T\}}(\mathbb{I}_{\{t_i<T_{(n)}\}} - e^{-k_d\int_t^{t_i} c(\xi_u)du})\Big]\Big|\,dx \to 0$$

as n tends to infinity. The first term of (2.16) is bounded above by a finite sum of such terms, therefore it converges to 0 $\mathbf{P}$-a.s, for each fixed $t \in [0,t_p]$.

Let $A_{\mathbf{f},\mathbf{t}}$ stand for the set

$$\Big\{(\omega,t)\in\mathbf{\Omega}\times[0,t_p] : \int_B dx\,\Big|\mathrm{E}_{t,x}\Big[\sum_{i=1}^p f_i(\xi_{t_i})\mathbb{I}_{\{t_i<T\}}(\mathbb{I}_{\{t_i<T_{(n)}\}} - e^{-k_d\int_t^{t_i} c(\xi_u)du})\Big]\Big| \to 0\Big\}.$$

If λ_1 denotes the Lebesgue measure on $\mathbb{R}$, we have by Fubini's theorem

$$\mathbf{P}\otimes\lambda_1(A^c_{\mathbf{f},\mathbf{t}}) = \int_0^{t_p} dt\,\mathbf{P}\big(\{\omega : (\omega,t)\in A^c_{\mathbf{f},\mathbf{t}}\}\big) = 0,$$

which gives (i) in the following lemma :

LEMMA 2.9. *(i) There exists a measurable subset $\tilde\Omega_{\mathbf{f},\mathbf{t}}$ of $\mathbf{\Omega}$, with $\mathbf{P}(\tilde\Omega_{\mathbf{f},\mathbf{t}}) = 0$, such that for every $\omega \in \mathbf{\Omega} \setminus \tilde\Omega_{\mathbf{f},\mathbf{t}}$,*

$$\int_B \left|\mathrm{E}_{t,x}\left[\sum_{i=1}^{p} f_i(\xi_{t_i})\mathbb{I}_{\{t_i<T\}}\left(\mathbb{I}_{\{t_i<T_{(n)}\}} - e^{-k_d\int_t^{t_i} c(\xi_u)du}\right)\right]\right| dx \to 0 \quad \text{as } n\to\infty$$

for all $t\geq 0$, except for t belonging to a Lebesgue null subset $\tilde T_{\mathbf{f},\mathbf{t},\omega}$ of $\mathbb{R}_+$.

(ii) There exists also a measurable subset $\hat\Omega_{\mathbf{f},\mathbf{t}}$ of $\mathbf{\Omega}$, with $\mathbf{P}(\hat\Omega_{\mathbf{f},\mathbf{t}}) = 0$, such that for every $\omega\in\mathbf{\Omega}\setminus\hat\Omega_{\mathbf{f},\mathbf{t}}$,

$$\int_0^\infty ds \int_B dx \left|\mathrm{E}_{0,x}\left[\mathbb{I}_{\{s<T\}}\left(\mathbb{I}_{\{s<T_{(n)}\}} - e^{-k_d\int_0^s c(\xi_u)du}\right)w_{s+t}^{(n)}(\xi_s)^2\right]\right| \to 0 \qquad \text{as } n\to\infty$$

for all $t\geq 0$, except on a Lebesgue null subset $\hat T_{\mathbf{f},\mathbf{t},\omega}$ of $\mathbb{R}_+$. Here, $w^{(n)} = w^{\varepsilon_n}$ is the function given by (2.13) corresponding to the superprocess $X^{(n)}$.

(iii) Finally, for all $x\in B$ there exists a negligible set $\Omega_{\mathbf{f},\mathbf{t},0}(x)$ outside which

$$\left|\mathrm{E}_{0,x}\left[\sum_{i=1}^{p} f_i(\xi_{t_i})\mathbb{I}_{\{t_i<T\}}\left(\mathbb{I}_{\{t_i<T_{(n)}\}} - e^{-k_d\int_0^{t_i} c(\xi_u)du}\right)\right]\right|$$

and

$$\left|\mathrm{E}_{0,x}\left[\int_0^\infty ds\, \mathbb{I}_{\{s<T\}}\left(\mathbb{I}_{\{s<T_{(n)}\}} - e^{-k_d\int_0^s c(\xi_u)du}\right)w_s^{(n)}(\xi_s)^2\right]\right|$$

converge to 0 as $n\to\infty$.

Both (ii) and (iii) can be obtained from Lemma 2.8 in a way similar to the derivation of (i). Note that in (ii), we may replace the integral over $[0,\infty)$ by the integral over $[0,t_p]$ (since $w_r^{(n)}\equiv 0$ if $r\geq t_p$) and that the functions $w_r^{(n)}$ are uniformly bounded by $C_{\mathbf{f}}$.

The first term in the right-hand side of (2.16), with $\varepsilon=\varepsilon_n$, converges to 0 as $n\to\infty$ provided that $\omega\notin\tilde\Omega_{\mathbf{f},\mathbf{t}}$ and $t\notin\tilde T_{\mathbf{f},\mathbf{t},\omega}$, by Lemma 2.9 (i). For the second term, we have

$$\begin{aligned}(2.18)\quad \int_B dx\,\left|\,\mathrm{E}_{t,x}\,\left[\int_t^\infty ds\,\mathbb{I}_{\{s<T\}}\left(\mathbb{I}_{\{s<T_{(n)}\}} - e^{-k_d\int_t^s c(\xi_u)du}\right)w_s^{(n)}(\xi_s)^2\right]\right| \\ = \int_B dx\,\left|\,\mathrm{E}_{0,x}\left[\int_0^\infty ds\,\mathbb{I}_{\{s<T\}}\left(\mathbb{I}_{\{s<T_{(n)}\}} - e^{-k_d\int_0^s c(\xi_u)du}\right)w_{s+t}^{(n)}(\xi_s)^2\right]\right| \\ \leq \int_0^\infty ds\,\int_B dx\,\left|\,\mathrm{E}_{0,x}\left[\mathbb{I}_{\{s<T\}}\left(\mathbb{I}_{\{s<T_{(n)}\}} - e^{-k_d\int_0^s c(\xi_u)du}\right)w_{s+t}^{(n)}(\xi_s)^2\right]\right|,\end{aligned}$$

which converges to 0 as $n\to\infty$ by Lemma 2.9 (ii), if $\omega\notin\hat\Omega_{\mathbf{f},\mathbf{t}}$ and $t\notin\hat T_{\mathbf{f},\mathbf{t},\omega}$.

Finally, for $\omega\in\left(\tilde\Omega_{\mathbf{f},\mathbf{t}}\cup\hat\Omega_{\mathbf{f},\mathbf{t}}\right)^c$ and $t\in\left(\tilde T_{\mathbf{f},\mathbf{t},\omega}\cup\hat T_{\mathbf{f},\mathbf{t},\omega}\right)^c$, the first two terms of the right-hand side of (2.16) converge to 0 as $n\to\infty$. Recalling (2.17), we obtain

$$(2.19)\qquad \int_B dx\,\left|w_t^{(n)}(x) - w_t^*(x)\right| \leq b_n(t) + 2C_{\mathbf{f}}\int_t^{t_p} ds\,\int_B dz\,\left|w_t^{(n)}(z) - w_t^*(z)\right|,$$

where $b_n(t)\to 0$ as $n\to\infty$ provided ω and t are as above. Besides, for every t,

$$(2.20)\qquad |b_n(t)| \leq 2\lambda(B)(1+t_p)(C_{\mathbf{f}} + C_{\mathbf{f}}^2).$$

Set for every $t\in[0,t_p]$,

$$G_n(t) := \int_B dx\,\left|w_{t_p-t}^{(n)}(x) - w_{t_p-t}^*(x)\right|.$$

Then,

$$G_n(t) \le b_n(t_p - t) + K_{\mathbf{f}} \int_0^t ds\, G_n(s),$$

where $K_{\mathbf{f}} := 2C_{\mathbf{f}}$. By iterating this inequality as in the proof of Gronwall's lemma, we obtain for all $k \ge 1$, $n \ge 1$ and $t \in [0, t_p]$

$$\begin{aligned} G_n(t) \le{}& b_n(t_p - t) + \sum_{i=0}^{k-2} K_{\mathbf{f}}^{i+1} \int_0^t ds_1 \int_0^{s_1} ds_2 \ldots \int_0^{s_i} ds_{i+1}\, b_n(t_p - s_{i+1}) \\ &+ K_{\mathbf{f}}^k \int_0^t ds_1 \int_0^{s_1} ds_2 \ldots \int_0^{s_{k-1}} ds_k G_n(s_k) \\ \le{}& b_n(t_p - t) + \sum_{i=0}^{k-2} K_{\mathbf{f}}^{i+1} \int_0^t ds_1 \int_0^{s_1} ds_2 \ldots \int_0^{s_i} ds_{i+1}\, b_n(t_p - s_{i+1}) + \lambda(B) K_{\mathbf{f}}^{k+1} \frac{t_p^k}{k!}\,. \end{aligned}$$

Fix $\varepsilon > 0$ and let $k \ge 2$ be such that $\lambda(B) K_{\mathbf{f}}^{k+1} \frac{t_p^k}{k!} \le \frac{\varepsilon}{2}$. For $\omega \notin \big(\tilde{\Omega}_{\mathbf{f},\mathbf{t}} \cup \hat{\Omega}_{\mathbf{f},\mathbf{t}}\big)$, $b_n(r)$ converges to 0 as $n \to \infty$ except on a Lebesgue null set of values of r, and thus by dominated convergence

$$\int_0^t ds_1 \int_0^{s_1} ds_2 \ldots \int_0^{s_i} ds_{i+1}\, b_n(t_p - s_{i+1}) \to 0$$

for every $t \in [0, t_p]$ and $i \in \{0, \ldots, k-2\}$. In particular, for such ω and for every $t \in [0, t_p]$, we have

$$K_{\mathbf{f}}^{i+1} \int_0^t ds_1 \int_0^{s_1} ds_2 \ldots \int_0^{s_i} ds_{i+1}\, b_n(t_p - s_{i+1}) \le \frac{\varepsilon}{2k}$$

for all n sufficiently large. If moreover t is such that $t_p - t \in \big(\tilde{T}_{\mathbf{f},\mathbf{t},\omega} \cup \hat{T}_{\mathbf{f},\mathbf{t},\omega}\big)^c$, then we have also $b_n(t_p - t) \le \frac{\varepsilon}{2k}$ if n is large. Hence, we have $G_n(t) \le \varepsilon$ when n is large. Since ε was arbitrary, we can conclude that for all ω and t as specified above, $G_n(t)$ converges to 0. Equivalently: for all $\omega \in \big(\tilde{\Omega}_{\mathbf{f},\mathbf{t}} \cup \hat{\Omega}_{\mathbf{f},\mathbf{t}}\big)^c$ and $t \in \big(\tilde{T}_{\mathbf{f},\mathbf{t},\omega} \cup \hat{T}_{\mathbf{f},\mathbf{t},\omega}\big)^c$,

$$\lim_{n \to \infty} \int_B dx\, \big|w_t^{(n)}(x) - w_t^*(x)\big| = 0. \tag{2.21}$$

We next consider the asymptotic behaviour of $|w_0^{(n)}(x) - w_0^*(x)|$. In the same way as in (2.16) but now without integrating over B, we have for every $x \in B$

$$\begin{aligned} \big|w_0^{(n)}(x) - w_0^*(x)\big| \le{}& \left| \mathrm{E}_{0,x}\left[\sum_{i=1}^p f_i(\xi_{t_i}) \mathbb{I}_{\{t_i < T\}} \Big(\mathbb{I}_{\{t_i < T_{(n)}\}} - e^{-k_d \int_0^{t_i} c(\xi_u) du} \Big) \right] \right| \\ &+ \left| \mathrm{E}_{0,x}\left[\int_0^\infty ds\, \mathbb{I}_{\{s < T\}} \Big(\mathbb{I}_{\{s < T_{(n)}\}} - e^{-k_d \int_0^s c(\xi_u) du} \Big) w_s^{(n)}(\xi_s)^2 \right] \right| \\ &+ 2C_{\mathbf{f}} \mathrm{E}_{0,x}\left[\int_0^\infty ds\, \mathbb{I}_B(\xi_s) \big| w_s^*(\xi_s) - w_s^{(n)}(\xi_s) \big| \right]. \end{aligned} \tag{2.22}$$

Let us fix $x \in B$. By Lemma 2.9 (iii), there exists a $\mathbf{P}$-negligible set $\Omega_{\mathbf{f},\mathbf{t},0}(x)$ outside which the first two terms in the right-hand side of (2.22) converge to 0. Besides, for any $\delta > 0$,

$$\begin{aligned} \mathrm{E}_{0,x}\left[\int_0^\infty ds\, \mathbb{I}_B(\xi_s) \big| w_s^*(\xi_s) - w_s^{(n)}(\xi_s) \big| \right] &= \int_0^\infty ds \int_B dz\, p_s(x,z) \big| w_s^*(z) - w_s^{(n)}(z) \big| \\ &\le 2C_{\mathbf{f}}\delta + \frac{1}{(2\pi\delta)^{d/2}} \int_\delta^\infty ds \int_B dz\, \big| w_s^*(z) - w_s^{(n)}(z) \big|, \end{aligned}$$

using the bound $p_s(x,z) \leq (2\pi\delta)^{-d/2}$ if $s \geq \delta$. If in addition $\omega \in \big(\tilde{\Omega}_{\mathbf{f},\mathbf{t}} \cup \hat{\Omega}_{\mathbf{f},\mathbf{t}}\big)^c$, then by (2.21) and dominated convergence (recall that w_s^* and $w_s^{(n)}$ vanish for $s > t_p$),

$$\int_\delta^\infty ds \int_B dz \,\big|w_s^*(z) - w_s^{(n)}(z)\big| \to 0$$

and so $\limsup |w_0^{(n)}(x) - w_0^*(x)| \leq 2C_{\mathbf{f}}\delta$. Since δ was arbitrary, it follows that $\lim |w_0^{(n)}(x) - w_0^*(x)| = 0$.

To summarize, for all $x \in B$ and $\omega \in \big(\tilde{\Omega}_{\mathbf{f},\mathbf{t}} \cup \hat{\Omega}_{\mathbf{f},\mathbf{t}} \cup \Omega_{\mathbf{f},\mathbf{t},0}(x)\big)^c$ (of $\mathbf{P}$-probability 1),

$$\lim_{n\to\infty} \big|w_0^{(n)}(x) - w_0^*(x)\big| = 0.$$

From the latter result, we can obtain the convergence of the finite-dimensional distributions of $X^{(n)}$ towards the corresponding ones for X^*. For all $x \in B$, $\mathbf{P}\big[w_0^{(n)}(x) \to w_0^*(x)\big] = 1$ so by applying once again Fubini's theorem, we have

$$\mathbf{P}\text{−a.s.},\quad \mu\text{−a.e.},\quad w_0^{(n)}(x) \to w_0^*(x) \quad \text{as } n \to \infty. \tag{2.23}$$

Since the $w^{(n)}$ are bounded by $C_{\mathbf{f}}$, dominated convergence and (2.23) give

$$\exp -\langle \mu, w_0^{(n)}(\cdot)\rangle \to \exp -\langle \mu, w_0^*(\cdot)\rangle.$$

Our construction from the historical superprocess makes it obvious that $X^{(n)}$ is stochastically bounded by X^0. It follows that the sequence of the distributions of $\{(X_{t_1}^{(n)}, \ldots, X_{t_p}^{(n)}), n \in \mathbb{N}\}$ is relatively compact. Therefore, if we choose a countable set of p-tuples $(f_1, \ldots, f_p)$ such that the corresponding family of maps $(\mu_1, \ldots, \mu_p) \mapsto \exp -\sum_{i=1}^p \langle \mu_i, f_i\rangle$ is convergence determining, we obtain that $(X_{t_1}^{(n)}, \ldots, X_{t_p}^{(n)})$ converges in distribution to $(X_{t_1}^*, \ldots, X_{t_p}^*)$ on a set of $\mathbf{P}$-probability 1 (which a priori depends on $(t_1, \ldots, t_p)$). This completes the proof of Proposition 2.6. □

Proof of Lemma 2.8. The quantity of interest vanishes if $t > t_1$, and so we need only consider the case $t \leq t_1$. In that case,

$$\begin{aligned}
&\mathbf{E}\Big[\Big(\mathrm{E}_{t,x}\Big[f(\xi_{t_1})\mathbb{I}_{\{t_1<T\}}\Big(\mathbb{I}_{\{t_1<T_\varepsilon\}} - e^{-k_d\int_t^{t_1} c(\xi_u)du}\Big)\Big]\Big)^2\Big] \\
&\quad = \mathbf{E}\Big[\mathrm{E}_{t,x}\Big[\mathbb{I}_{\{t_1<T\}}\mathbb{I}_{\{t_1<T'\}} f(\xi_{t_1}) f(\xi'_{t_1})\Big(\mathbb{I}_{\{t_1<T_\varepsilon\}}\mathbb{I}_{\{t_1<T'_\varepsilon\}} - \mathbb{I}_{\{t_1<T_\varepsilon\}} e^{-k_d\int_t^{t_1} c(\xi'_u)du} \\
&\qquad - \mathbb{I}_{\{t_1<T'_\varepsilon\}} e^{-k_d\int_t^{t_1} c(\xi_u)du} + e^{-k_d\int_t^{t_1}(c(\xi_u)+c(\xi'_u))du}\Big)\Big]\Big],
\end{aligned} \tag{2.24}$$

where ξ' is another Brownian motion, independent of ξ, T' and T'_ε are defined in an obvious way and we have kept the notation $\mathrm{P}_{t,x}$ for the probability measure under which the two Brownian motions start from x at time t. Recall that $S_\varepsilon(s,t)$ denotes the Wiener sausage of radius ε along the time interval $[s,t]$ associated to ξ, and define $S'_\varepsilon(s,t)$ in a similar way. Then,

$$\begin{aligned}
\mathbf{E}\big[\mathbb{I}_{\{t_1<T_\varepsilon\}}\mathbb{I}_{\{t_1<T'_\varepsilon\}}\big] &= \mathbf{P}\big[\mathcal{P}^\varepsilon \cap (S_\varepsilon(t,t_1) \cup S'_\varepsilon(t,t_1)) = \emptyset\big] \\
&= \exp\Big\{ - s_d(\varepsilon)\int_{S_\varepsilon(t,t_1)\cup S'_\varepsilon(t,t_1)} c(y)dy\Big\}
\end{aligned}$$

and similarly

$$\mathbf{E}\big[\mathbb{I}_{\{t_1<T_\varepsilon\}}\big] = \exp\Big\{ - s_d(\varepsilon)\int_{S_\varepsilon(t,t_1)} c(y)dy\Big\}.$$

Set $\tilde{t}_1 = t_1 - t$. By Fubini's theorem and a simple symmetry argument, the quantity in (2.24) is equal to

$$\begin{aligned}
&\mathrm{E}_{t,x}\Big[\mathbb{I}_{\{t_1<T\}}\,\mathbb{I}_{\{t_1<T'\}}f(\xi_{t_1})f(\xi'_{t_1})\Big\{2e^{-k_d\int_t^{t_1}c(\xi'_u)du}\big(e^{-k_d\int_t^{t_1}c(\xi'_u)du}-e^{-s_d(\varepsilon)\int_{S_\varepsilon(t,t_1)}c(y)dy}\big)\\
&\qquad\qquad +e^{-s_d(\varepsilon)\int_{S_\varepsilon(t,t_1)\cup S'_\varepsilon(t,t_1)}c(y)dy}-e^{-k_d\int_t^{t_1}(c(\xi_u)+c(\xi'_u))du}\Big\}\Big]\\
\leq\quad & 2\|f\|^2\mathrm{E}_{t,x}\Big[\Big|e^{-k_d\int_t^{t_1}c(\xi'_u)du}-e^{-s_d(\varepsilon)\int_{S_\varepsilon(t,t_1)}c(y)dy}\Big|\Big]\\
&+\|f\|^2\mathrm{E}_{t,x}\Big[\Big|e^{-s_d(\varepsilon)\int_{S_\varepsilon(t,t_1)\cup S'_\varepsilon(t,t_1)}c(y)dy}-e^{-k_d\int_t^{t_1}(c(\xi_u)+c(\xi'_u))du}\Big|\Big]\\
\leq\quad & 2\|f\|^2\mathrm{E}_{0,x}\Big[\Big|\,k_d\int_0^{\tilde{t}_1}c(\xi_u)du-s_d(\varepsilon)\int_{S_\varepsilon(0,\tilde{t}_1)}c(y)dy\Big|\Big]\\
&+\|f\|^2\mathrm{E}_{0,x}\Big[\Big|s_d(\varepsilon)\int_{S_\varepsilon(0,\tilde{t}_1)}c(y)dy+s_d(\varepsilon)\int_{S'_\varepsilon(0,\tilde{t}_1)}c(y)dy\\
&-s_d(\varepsilon)\int_{S_\varepsilon(0,\tilde{t}_1)\cap S'_\varepsilon(0,\tilde{t}_1)}c(y)dy-k_d\int_0^{\tilde{t}_1}c(\xi_u)du-k_d\int_0^{\tilde{t}_1}c(\xi'_u)du\Big|\Big]\\
\leq\quad & 4\|f\|^2\mathrm{E}_{0,x}\Big[\Big|\,k_d\int_0^{\tilde{t}_1}c(\xi_u)du-s_d(\varepsilon)\int_{S_\varepsilon(0,\tilde{t}_1)}c(y)dy\Big|\Big]\\
&+\|f\|^2\|c\|s_d(\varepsilon)\mathrm{E}_{0,x}\big[\lambda\big(S_\varepsilon(0,\tilde{t}_1)\cap S'_\varepsilon(0,\tilde{t}_1)\big)\big],
\end{aligned}\tag{2.25}$$

where in the second inequality we used the bound $|e^{-x}-e^{-y}|\leq|x-y|$ for $x,y\geq 0$.

On the one hand, by **[LG86]** ($d=2,3$) and **[LG88]** p.1009-1010 ($d\geq 4$), we have

$$\begin{aligned}
|\log\varepsilon|\;\mathrm{E}_{0,x}\big[\lambda\big(S_\varepsilon(0,\tilde{t}_1)\cap S'_\varepsilon(0,\tilde{t}_1)\big)\big] &= \mathcal{O}(|\log\varepsilon|^{-1}) && \text{if } d=2,\\
\varepsilon^{-1}\mathrm{E}_{0,x}\big[\lambda\big(S_\varepsilon(0,\tilde{t}_1)\cap S'_\varepsilon(0,\tilde{t}_1)\big)\big] &= \mathcal{O}(\varepsilon) && \text{if } d=3,\\
\varepsilon^{2-d}\mathrm{E}_{0,x}\big[\lambda\big(S_\varepsilon(0,\tilde{t}_1)\cap S'_\varepsilon(0,\tilde{t}_1)\big)\big] &= \mathcal{O}(\varepsilon^2|\log\varepsilon|) && \text{if } d\geq 4.
\end{aligned}\tag{2.26}$$

On the other hand,

$$\begin{aligned}
&\mathrm{E}_{0,x}\Big[\Big|\,k_d\int_0^{\tilde{t}_1}c(\xi_u)du-s_d(\varepsilon)\int_{S_\varepsilon(0,\tilde{t}_1)}c(y)dy\Big|\Big]\\
&\qquad\qquad\leq \mathrm{E}_{0,x}\Big[\Big(k_d\int_0^{\tilde{t}_1}c(\xi_u)du-s_d(\varepsilon)\int_{S_\varepsilon(0,\tilde{t}_1)}c(y)dy\Big)^2\Big]^{1/2}.
\end{aligned}\tag{2.27}$$

Proposition 2.3 ensures that the right-hand side of (2.27) is bounded by $K\|c\|\;|\log\varepsilon|^{-1}$ if $d=2$ and by $K\|c\|\;\varepsilon|\log\varepsilon|$ if $d\geq 3$. Together with (2.25) and (2.26), this completes the proof of Lemma 2.8. ☐

4. Tightness of the sequence $X^{(n),B}$

Let $C^2_+(\mathbb{R}^d)$ denote the set of all nonnegative functions of class C^2 on $\mathbb{R}^d$. By Theorem II.4.1 in [**Per02**], the tightness of the sequence of the laws of the superprocesses $X^{(n),B}$ will follow if we can prove that the sequence of the laws of $\langle X^{(n),B},\phi\rangle$ is tight, for every $\phi\in C^2_+(\mathbb{R}^d)$ with compact support. Note that condition (i) in Theorem II.4.1 of [**Per02**] holds thanks to the domination $X^{(n),B}\leq X^0$. Recall that X^0 is the usual super-Brownian motion without obstacles.

Let us first introduce the $\mathbf{P}$-negligible set $\Theta\subset\mathbf{\Omega}$ outside which the desired tightness will hold.

DEFINITION 2.10. ***(Good environments)*** *Let Θ be the union over all choices of the integer $p\geq 1$ and of the rational numbers $q_1,\dots,q_p$ of the $\mathbf{P}$-negligible sets on which the sequence*

$(X^{(n)}_{q_1}, \ldots, X^{(n)}_{q_p})$ *does not converge in distribution to* $(X^*_{q_1}, \ldots, X^*_{q_p})$ *as* $n \to \infty$. *We call* good environment *any environment which does not belong to* Θ.

To simplify notation, we again write $X^{(n)}$ for $X^{(n),B}$ (as in the last section, B is fixed) and prove tightness only on the time interval $[0,1]$. Let us fix $\phi \in C^2_+(\mathbb{R}^d)$ with compact support. The tightness of the sequence $\langle X^{(n)}, \phi\rangle$ is a consequence of the following lemma.

LEMMA 2.11. *If* $\omega \notin \Theta$, *then for every* $\varepsilon > 0$ *there exists* $k = k(\varepsilon) \geq 1$ *and* $n_0 = n_0(\omega, \varepsilon, k)$ *such that for all* $n \geq n_0$,

$$\mathbb{P}_\mu\left[\bigcup_{i=0}^{k-1}\left\{\sup_{\frac{i}{k}\leq t\leq\frac{i+1}{k}} \left|\langle X^{(n)}_t, \phi\rangle - \langle X^{(n)}_{\frac{i}{k}}, \phi\rangle\right| > \varepsilon\right\}\right] < \varepsilon.$$

Lemma 2.11 easily implies that the sequence $\langle X^{(n)}, \phi\rangle$ is tight. Indeed, let us fix a good environment and $\eta > 0$. By Lemma 2.11, there exist $k(\eta)$ and $n_0(\omega, \eta, k)$ such that for all $n \geq n_0$,

$$\mathbb{P}_\mu\left[\bigcup_{i=0}^{k-1}\left\{\sup_{\frac{i}{k}\leq t\leq\frac{i+1}{k}} \left|\langle X^{(n)}_t, \phi\rangle - \langle X^{(n)}_{\frac{i}{k}}, \phi\rangle\right| > \frac{\eta}{3}\right\}\right] < \eta. \tag{2.28}$$

If $n \geq n_0$ is fixed, on the complement of the event considered in (2.28), we have for every $s, t \in [0,1]$

$$|t-s| \leq \frac{1}{k} \quad \Rightarrow \quad \left|\langle X^{(n)}_t, \phi\rangle - \langle X^{(n)}_s, \phi\rangle\right| \leq \eta$$

and therefore $w'(\langle X^{(n)}, \phi\rangle, \frac{1}{k}, 1) \leq \eta$, using the notation of Ethier and Kurtz [**EK86**] for the modulus of continuity of the process $\langle X^{(n)}, \phi\rangle$. Thus, for all $n \geq n_0$,

$$\mathbb{P}_\mu\Big[w'\big(\langle X^{(n)}, \phi\rangle, \frac{1}{k}, 1\big) \leq \eta\Big] \geq \mathbb{P}_\mu\left[\left(\bigcup_{i=0}^{k-1}\left\{\sup_{\frac{i}{k}\leq t\leq\frac{i+1}{k}} \left|\langle X^{(n)}_t, \phi\rangle - \langle X^{(n)}_{\frac{i}{k}}, \phi\rangle\right| > \frac{\eta}{3}\right\}\right)^c\right] \geq 1-\eta.$$

In addition, ϕ is bounded so the first condition of Theorem 3.7.2 in [**EK86**] is trivially fulfilled, hence Corollary 3.7.4 of [**EK86**] implies that for any good environment, the sequence of the laws of $\langle X^{(n)}, \phi\rangle$ under $\mathbb{P}_\mu$ is tight.

Let us now turn to the proof of Lemma 2.11.

Proof of Lemma 2.11. We fix a good environment. Let $\varepsilon > 0$. The process $(\langle X^*_t, \phi\rangle)_{t\geq 0}$ is continuous, therefore there exists $k_0(\varepsilon)$ such that for all $k \geq k_0$,

$$\mathbb{P}_\mu\left[\sup_{0\leq i\leq k-1} \left|\left\langle X^*_{\frac{i+1}{k}}, \phi\right\rangle - \left\langle X^*_{\frac{i}{k}}, \phi\right\rangle\right| \geq \frac{\varepsilon}{2}\right] < \frac{\varepsilon}{3}. \tag{2.29}$$

There exists $K = K(\varepsilon) \geq 1$ such that

$$\mathbb{P}_\mu\left[\sup_{0\leq t\leq 1} \langle X^0_t, 1\rangle \geq K\right] < \frac{\varepsilon}{3}. \tag{2.30}$$

By a trivial domination argument, the bound (2.30) remains valid if we replace X^0 by $X^{(n)}$ (in fact for any environment). In the following, we fix the constant $K \geq 1$ such that (2.30) holds.

We now have the following result :

LEMMA 2.12. *There exists a constant* $C = C(\phi, K)$ *such that for every integer* $k \geq 1$ *and every measure* $\gamma \in \mathcal{M}_f(\mathbb{R}^d)$ *satisfying* $\langle \gamma, 1\rangle \leq K$,

$$\mathbb{P}_\gamma\left[\sup_{0\leq s\leq\frac{1}{k}} \left|\langle X^0_s, \phi\rangle - \langle X^0_0, \phi\rangle\right| > \frac{\varepsilon}{2}\right] \leq \frac{C}{k^2}\,.$$

The proof of Lemma 2.12 is deferred to the end of the section. Let us define

$$A_n = \Big\{ \sup_{0\le t\le 1} \langle X_t^{(n)}, 1\rangle \ge K \Big\}.$$

Then,

$$\begin{aligned}
&\mathbb{P}_\mu\left[\bigcup_{i=0}^{k-1} \Big\{ \sup_{\frac{i}{k}\le t\le \frac{i+1}{k}} \big|\langle X_t^{(n)},\phi\rangle - \langle X_{\frac{i}{k}}^{(n)},\phi\rangle\big| > \varepsilon \Big\}\right] \\
&\quad \le \mathbb{P}_\mu[A_n] + \mathbb{P}_\mu\left[\sup_{0\le i\le k-1} \Big|\langle X_{\frac{i+1}{k}}^{(n)},\phi\rangle - \langle X_{\frac{i}{k}}^{(n)},\phi\rangle\Big| > \frac{\varepsilon}{2}\right] \\
&\quad +\mathbb{P}_\mu\left[A_n^c \cap \Big\{ \sup_{0\le i\le k-1} \Big\{ \sup_{\frac{i}{k}\le t\le\frac{i+1}{k}} \Big(\langle X_t^{(n)},\phi\rangle - \langle X_{\frac{i}{k}}^{(n)},\phi\rangle\Big)\Big\} > \varepsilon\Big\}\right] \\
&\quad +\mathbb{P}_\mu\left[A_n^c \cap \Big\{ \sup_{0\le i\le k-1} \Big\{ \sup_{\frac{i}{k}\le t\le\frac{i+1}{k}} \Big(\langle X_{\frac{i}{k}}^{(n)},\phi\rangle - \langle X_t^{(n)},\phi\rangle\Big)\Big\} > \varepsilon\Big\} \cap \right. \\
&\qquad\qquad \left. \Big\{ \sup_{0\le i\le k-1} \Big|\langle X_{\frac{i+1}{k}}^{(n)},\phi\rangle - \langle X_{\frac{i}{k}}^{(n)},\phi\rangle\Big| \le \frac{\varepsilon}{2}\Big\}\right] \\
&\quad = a_n + b_n + c_n + d_n
\end{aligned}$$

From (2.30), we have

$$a_n < \frac{\varepsilon}{3}.$$

Moreover, from the definition of a good environment, if $k\ge k_0$,

$$\limsup_{n\to\infty} b_n \le \mathbb{P}_\mu\left[\sup_{0\le i\le k-1} \Big|\langle X^*_{\frac{i+1}{k}},\phi\rangle - \langle X^*_{\frac{i}{k}},\phi\rangle\Big| \ge \frac{\varepsilon}{2}\right] < \frac{\varepsilon}{3},$$

by (2.29). Thus if $k\ge k_0$, there exists $n_0(\omega,\varepsilon,k)$ such that for all $n\ge n_0$, $b_n(k)\le \frac{\varepsilon}{3}$. Then,

$$\begin{aligned}
c_n &\le \sum_{i=0}^{k-1} \mathbb{P}_\mu\left[\langle X_{i/k}^{(n)},1\rangle \le K;\ \sup_{\frac{i}{k}\le t\le\frac{i+1}{k}} \Big(\langle X_t^{(n)},\phi\rangle - \langle X_{\frac{i}{k}}^{(n)},\phi\rangle\Big) > \varepsilon\right] \\
&= \sum_{i=0}^{k-1} \mathbb{E}_\mu\left[\mathbb{I}_{\{\langle X_{i/k}^{(n)},1\rangle\le K\}}\, \mathbb{P}_{X_{i/k}^{(n)}}\left[\sup_{0\le t\le\frac{1}{k}} \big(\langle X_t^{(n)},\phi\rangle - \langle X_0^{(n)},\phi\rangle\big) > \varepsilon\right]\right].
\end{aligned}$$

The last equality is obtained by applying the Markov property to $X^{(n)}$ at time $\frac{i}{k}$. By a domination argument, we have for all $\gamma\in\mathcal{M}_f(\mathbb{R}^d)$ such that $\langle\gamma,1\rangle\le K$,

$$\mathbb{P}_\gamma\left[\sup_{0\le t\le\frac{1}{k}} \big(\langle X_t^{(n)},\phi\rangle - \langle X_0^{(n)},\phi\rangle\big) > \varepsilon\right] \le \mathbb{P}_\gamma\left[\sup_{0\le t\le\frac{1}{k}} \big(\langle X_t^{0},\phi\rangle - \langle X_0^{0},\phi\rangle\big) > \varepsilon\right] \le \frac{C}{k^2}$$

by Lemma 2.12. It follows that

$$c_n \le k.\frac{C}{k^2} = \frac{C}{k} < \frac{\varepsilon}{6}$$

if $k\ge k_1(\varepsilon)$. Finally,

$$\begin{aligned}
d_n \le \sum_{i=0}^{k-1} \mathbb{P}_\mu\Bigg[&\sup_{\frac{i}{k}\le t\le\frac{i+1}{k}} \langle X_t^{(n)},1\rangle \le K;\ \sup_{\frac{i}{k}\le t\le\frac{i+1}{k}} \Big(\langle X_{\frac{i}{k}}^{(n)},\phi\rangle - \langle X_t^{(n)},\phi\rangle\Big) > \varepsilon; \\
&\sup_{0\le i\le k-1} \Big|\langle X_{\frac{i+1}{k}}^{(n)},\phi\rangle - \langle X_{\frac{i}{k}}^{(n)},\phi\rangle\Big| \le \frac{\varepsilon}{2}\Bigg].
\end{aligned}$$

We fix $i \in \{0, \ldots, k-1\}$ and consider the stopping time

$$T_i := \inf \Big\{ t \geq \frac{i}{k} : \ \langle X_t^{(n)}, \phi \rangle \leq \langle X_{\frac{i}{k}}^{(n)}, \phi \rangle - \varepsilon \Big\}.$$

Then, the i-th term of the previous sum is bounded by

$$\mathbb{P}_\mu \left[T_i \leq \frac{i+1}{k}, \ \langle X_{T_i}^{(n)}, 1 \rangle \leq K, \sup_{T_i \leq t \leq T_i + \frac{1}{k}} \Big(\langle X_t^{(n)}, \phi \rangle - \langle X_{T_i}^{(n)}, \phi \rangle \Big) \geq \frac{\varepsilon}{2} \right]$$

$$= \mathbb{E}_\mu \Bigg[\mathbb{I}_{\left\{ T_i \leq \frac{i+1}{k}, \ \langle X_{T_i}^{(n)}, 1 \rangle \leq K \right\}} \mathbb{P}_{X_{T_i}^{(n)}} \Bigg[\sup_{0 \leq t \leq \frac{1}{k}} \left(\langle X_t^{(n)}, \phi \rangle - \langle X_0^{(n)}, \phi \rangle \right) \geq \frac{\varepsilon}{2} \Bigg] \Bigg]$$

by the strong Markov property at time T_i. Using Lemma 2.12 once again, we see that this quantity is bounded by C/k^2, hence for $k \geq k_1(\varepsilon)$,

$$d_n \leq \frac{C}{k} \leq \frac{\varepsilon}{6}.$$

Combining the preceding estimates, we obtain that for $k = k_0(\varepsilon) \vee k_1(\varepsilon)$, and every $n \geq n_0(\omega, \varepsilon, k)$,

$$\mathbb{P}_\mu \Bigg[\bigcup_{i=0}^{k-1} \Big\{ \sup_{\frac{i}{k} \leq t \leq \frac{i+1}{k}} \left| \langle X_t^{(n)}, \phi \rangle - \langle X_{\frac{i}{k}}^{(n)}, \phi \rangle \right| > \varepsilon \Big\} \Bigg] < \varepsilon.$$

This completes the proof of Lemma 2.11. □

Proof of Lemma 2.12. Let $\gamma \in \mathcal{M}_f(\mathbb{R}^d)$ be such that $|\gamma| := \langle \gamma, 1 \rangle \leq K$. Recall that the process $(\langle X_t^0, 1 \rangle)_{t \geq 0}$ is a martingale. From the maximal inequality applied to the nonnegative submartingale $(\langle X_t^0, 1 \rangle - |\gamma|)^4$,

$$\mathbb{P}_\gamma \Bigg[\sup_{0 \leq t \leq \frac{1}{k}} \langle X_t^0, 1 \rangle > 2K \Bigg] \leq \frac{1}{(2K - |\gamma|)^4} \, \mathbb{E}_\gamma \Big[\big(\langle X_{\frac{1}{k}}^0, 1 \rangle - |\gamma| \big)^4 \Big].$$

We claim that

$$\mathbb{E}_\gamma \Big[\big(\langle X_{\frac{1}{k}}^0, 1 \rangle - |\gamma| \big)^4 \Big] = \frac{24}{k^3} |\gamma| + \frac{12}{k^2} |\gamma|^2. \tag{2.31}$$

To prove this claim, recall that $Y_t = \langle X_t^0, 1 \rangle$ is a Feller diffusion, whose semigroup Laplace transform is given by

$$\mathbb{E}\big[\exp -\lambda Y_t | \, Y_0 = y \big] = \exp \Big(-\frac{\lambda y}{1 + \lambda t} \Big),$$

for $\lambda \geq 0$. Thus,

$$\mathbb{E}\big[\exp -\lambda (Y_t - y) | \, Y_0 = y \big] = \exp \Big(\frac{\lambda^2 t y}{1 + \lambda t} \Big) = 1 + \lambda^2 t y - \lambda^3 t^2 y + \lambda^4 t^3 y + \frac{\lambda^4 t^2 y^2}{2} + o(\lambda^4),$$

as $\lambda \to 0$. From this expansion of the Laplace transform, we derive that

$$\mathbb{E}\Big[\big(Y_t - y \big)^4 | Y_0 = y \Big] = 24 t^3 y + 12 t^2 y^2,$$

which proves our claim (2.31). It follows that

$$\mathbb{P}_\gamma \Bigg[\sup_{0 \leq t \leq \frac{1}{k}} \langle X_t^0, 1 \rangle > 2K \Bigg] \leq \frac{12 |\gamma| (|\gamma| + 2)}{(2K - |\gamma|)^4 \, k^2}.$$

Let $A_{K,k}$ be the event $\big\{ \sup_{0 \leq t \leq \frac{1}{k}} \langle X_t^0, 1 \rangle > 2K \big\}$ and $B_{K,k}$ the event $\big\{ \sup_{0 \leq t \leq \frac{1}{k}} \big| \langle X_t^0, \phi \rangle - \langle X_0^0, \phi \rangle \big| > \frac{\varepsilon}{2} \big\}$. Then,

$$\mathbb{P}_\gamma [B_{K,k}] \leq \mathbb{P}_\gamma [A_{K,k}] + \mathbb{P}_\gamma [A_{K,k}^c \cap B_{K,k}] \leq \frac{c_0}{k^2} + \mathbb{P}_\gamma [A_{K,k}^c \cap B_{K,k}], \tag{2.32}$$

where c_0 is a constant depending on K.

In addition, $M_t := \langle X^0_t, \phi\rangle - \langle X^0_0, \phi\rangle - \int_0^t dr\ \langle X^0_r, \frac{1}{2}\Delta\phi\rangle$ is a continuous martingale with quadratic variation $2\int_0^t dr\ \langle X^0_r, \phi^2\rangle$. By the Dubins-Schwarz theorem (see Theorem V.1.7 in [**RY99**]), there exists a standard one-dimensional Brownian motion W such that $M_t = W_{\langle M\rangle_t}$ for all $t \geq 0$ a.s. On the event $A^c_{K,k}$, we have

$$\Big|\int_0^t dr\ \langle X^0_r, \frac{1}{2}\Delta\phi\rangle\Big| \leq t\|\Delta\phi\|K \leq \frac{c_1}{k} \quad \text{if } t \in [0, k^{-1}]$$

and

$$\langle M\rangle_t \leq \frac{4\|\phi\|^2 K}{k} = \frac{c_2}{k} \quad \text{if } t \in [0, k^{-1}],$$

where c_1 and c_2 are constants depending on ϕ and on K. Choose k_0 such that $c_1 k^{-1} < \varepsilon/4$ for every $k \geq k_0$. Then for all $k \geq k_0$,

$$\begin{aligned}
\mathbb{P}_\gamma\big[A^c_{K,k} \cap B_{K,k}\big] &= \mathbb{P}_\gamma\left[A^c_{K,k} \cap \left\{\sup_{0\leq t\leq 1/k}\left|M_t + \int_0^t dr\ \left\langle X^0_r, \frac{1}{2}\Delta\phi\right\rangle\right| > \frac{\varepsilon}{2}\right\}\right]\\
&\leq \mathbb{P}_\gamma\left[A^c_{K,k} \cap \left\{\sup_{0\leq t\leq 1/k}|M_t| > \frac{\varepsilon}{4}\right\}\right]\\
&\leq \mathrm{P}\left[\sup_{0\leq t\leq (c_2/k)}|W_t| > \frac{\varepsilon}{4}\right]\\
&\leq \frac{c_3}{k^2}\ ,
\end{aligned}$$

where c_3 is a constant depending on ϕ, K and ε. Together with (2.32), this completes the proof of Lemma 2.12. □

5. Proof of Theorem 2.1 and Corollary 2.2

The proof of Theorem 2.1 in the case when D is bounded is easy from the results of the previous sections. Let us take $B = D$ and recall the notation Θ for the union of the $\mathbf{P}$-negligible set on which there exist rational numbers $t_1, \ldots, t_p$ such that $(X^{(n)}_{t_1}, \ldots, X^{(n)}_{t_p})$ does not converge to $(X^*_{t_1}, \ldots, X^*_{t_p})$. The set Θ is $\mathbf{P}$-negligible and we showed in the previous section that the sequence $X^{(n),B}$ is tight on Θ^c. Hence, Theorem 3.7.8 of [**EK86**] allows us to conclude that for every $\omega \in \Theta^c$, $X^{(n)} \overset{(d)}{\to} X^*$ as $n \to \infty$.

We can now use the previous result to complete the proof of Theorem 2.1 when D is a domain of $\mathbb{R}^d$ which is not necessarily bounded.

Proof of Theorem 2.1 for a general domain D. Let μ be a finite measure on D and suppose first that the support of μ is bounded. Under $\mathbb{P}_\mu$, the superprocesses $X^{(n)}$ are stochastically dominated by the superprocess X^0, whose range

$$\mathcal{R}(X^0) = \overline{\bigcup_{t\geq\, 0} \mathrm{supp} X^0_t}$$

is almost surely compact since its initial value has compact support. Consequently, for every $\varepsilon > 0$, there exists a bounded open subset B of D containing the support of μ such that, for every environment and every $n \geq 1$,

$$\mathbb{P}_\mu\big[\mathcal{R}(X^{(n),D}) \subset B\big] \geq 1 - \varepsilon$$

and

$$\mathbb{P}_\mu\big[\mathcal{R}(X^{*,D}) \subset B\big] \geq 1 - \varepsilon.$$

From these inequalities, we can deduce that

$$d(\mathbb{P}_\mu^{(n),D},\mathbb{P}_\mu^{(n),B}) \le 2\varepsilon, \qquad n \ge 1; \tag{2.33}$$

$$d(\mathbb{P}_\mu^{*,D},\mathbb{P}_\mu^{*,B}) \le 2\varepsilon, \tag{2.34}$$

where d is the Prohorov metric on $\mathcal{M}_1(D_{\mathcal{M}_f(D)}[0,\infty))$. By the results of the last two sections, with $\mathbf{P}$-probability 1 there exists an integer $n_0(\omega)$ such that for all $n \ge n_0$,

$$d(\mathbb{P}_\mu^{(n),B},\mathbb{P}_\mu^{*,B}) \le \varepsilon.$$

Together with (2.33) and (2.34), this yields

$$d(\mathbb{P}_\mu^{(n),D},\mathbb{P}_\mu^{*,D}) \le 5\varepsilon \qquad\qquad \text{for all } n \ge n_0(\omega),$$

hence we can conclude that $\mathbb{P}_\mu^{(n),D}$ converges towards $\mathbb{P}_\mu^{*,D}$ on a set of $\mathbf{P}$-probability 1.

Finally, if the support of μ is unbounded, we can replace μ by the measure $\tilde\mu$ defined as the restriction of μ to a large ball centered at the origin. Using once again the domination of $X^{(n),D}$ (for all $n \ge 1$) and of $X^{*,D}$ by X^0, the law of $X^{(n),D}$ under $\mathbb{P}_\mu$ can be approximated uniformly in n by the law of $X^{(n),D}$ under $\mathbb{P}_{\tilde\mu}$, and similarly for $X^{*,D}$. The desired result then follows from the bounded support case. We leave the details to the reader. □

We end this section with the proof of Corollary 2.2.

Proof of Corollary 2.2. Let us argue by contradiction and suppose that there exist $\delta > 0$ and a sequence $\{\varepsilon_k, k \in \mathbb{N}\}$ decreasing to zero such that for all $k \ge 1$,

$$\mathbf{P}\left[d\big(\mathbb{P}_\mu^{\varepsilon_k,D},\mathbb{P}_\mu^{*,D}\big) > \delta\right] \ge \delta. \tag{2.35}$$

By extracting a subsequence, we can always choose ε_k such that

$$\sum_{k=1}^{\infty} |\log\varepsilon_k|^{-1} < \infty \;\text{ if } d = 2,$$

or

$$\sum_{k=1}^{\infty} \varepsilon_k\,|\log\varepsilon_k| < \infty \;\text{ if } d \ge 3.$$

But the latter condition is the only requirement for the sequence of superprocesses $X^{\varepsilon_k,D}$ to converge in distribution to $X^{*,D}$ with $\mathbf{P}$-probability 1, yielding a contradiction with (2.35). □

6. Appendix: Proof of (2.11)

The bound (2.11) is a consequence of the following lemma.

LEMMA 2.13. *There exists a function $\varphi : \mathbb{R}^2 \to [0,\infty]$ such that $\int_{\mathbb{R}^2}\varphi(y)dy < \infty$ and for every $y \in \mathbb{R}^2$ and $\varepsilon \in (0,\frac12)$,*

$$\left|\mathrm{P}_0\big[y \in S_\varepsilon(0,1)\big] - \frac{\pi}{|\log\varepsilon|}\int_0^1 ds\, p_s(y)\right| \le \frac{\varphi(y)}{|\log\varepsilon|^2},$$

where $p_s(y) = (2\pi s)^{-1}\exp\big\{-|y|^2/(2s)\big\}$ is the Brownian transition density.

REMARK 2.14. *The convergence of $|\log\varepsilon|\mathrm{P}_0[y \in S_\varepsilon(0,1)]$ towards $\pi\int_0^1 ds\, p_s(y)$ as ε tends to 0 was first obtained by Spitzer* **[Spi58]**. *See also* **[LG86]** *and* **[Szn87]** *for related results.*

Before proving Lemma 2.13, let us use it to derive (2.11). If c is a bounded nonnegative measurable function on $\mathbb{R}^2$ such that $\|c\| \leq 1$, then for every $\varepsilon \in (0, \frac{1}{2}]$ and $z \in \mathbb{R}^2$,

$$\begin{aligned}\Big|\mathrm{E}_z\Big[|\log\varepsilon| \int_{S_\varepsilon(0,1)} c(y)\,dy - \pi\int_0^1 c(\xi_s)ds\Big]\Big| \\ &= \Big|\int dy\, c(z+y)\Big(|\log\varepsilon|\,\mathrm{P}_0\big[y \in S_\varepsilon(0,1)\big] - \pi\int_0^1 ds\, p_s(y)\Big)\Big| \\ &\leq\ |\log\varepsilon|^{-1}\int \varphi(y)dy,\end{aligned}$$

which is the desired result.

Let us hence establish Lemma 2.13. The following proof is due to J.-F. Le Gall (personal communication).

Proof of Lemma 2.13. If $|y| \leq 10\varepsilon$, simple estimates show that the bound of Lemma 2.13 holds with $\varphi(y) = C\big((\log|y|)^2 + 1\big)$ for a suitable constant C. So we assume that $|y| > 10\varepsilon$. We put

$$\tau_\varepsilon(y) = \inf\big\{t \geq 0 : |\xi_t - y| \leq \varepsilon\big\}$$

in such a way that $\{y \in S_\varepsilon(0,1)\} = \{\tau_\varepsilon(y) \leq 1\}$. Let a_ε be an arbitrary point of the circle of radius ε centered at the origin, and

$$f(\varepsilon) = \mathrm{E}_{a_\varepsilon}\Big[\int_0^1 ds\, \mathbb{I}_{\{|\xi_s| \leq \varepsilon\}}\Big].$$

A straightforward calculation gives

$$f(\varepsilon) = \varepsilon^2|\log\varepsilon| + \mathcal{O}(\varepsilon^2) \tag{2.36}$$

as $\varepsilon \longrightarrow 0$.

Lower bound. An application of the strong Markov property at time $\tau_\varepsilon(y)$ shows for every $u \in (0,1]$, that

$$\mathrm{E}_0\Big[\int_0^u ds\, \mathbb{I}_{\{|\xi_s - y| \leq \varepsilon\}}\Big] \leq \mathrm{P}_0\big[\tau_\varepsilon(y) \leq u\big] f(\varepsilon).$$

On the other hand,

$$\mathrm{E}_0\Big[\int_0^u ds\, \mathbb{I}_{\{|\xi_s - y| \leq \varepsilon\}}\Big] = \int_0^u ds \int_{|z-y|\leq\varepsilon} dz\, p_s(z),$$

and thus

$$\begin{aligned}\Big|\mathrm{E}_0\Big[\int_0^u ds\, \mathbb{I}_{\{|\xi_s - y| \leq \varepsilon\}}\Big] - \pi\varepsilon^2\int_0^u ds\, p_s(y)\Big| \\ &\leq \int_0^u \frac{ds}{2\pi s}\int_{|z-y|\leq\varepsilon} dz\, \Big|\exp\Big\{-\frac{|z|^2}{2s}\Big\} - \exp\Big\{-\frac{|y|^2}{2s}\Big\}\Big| \\ &\leq \int_0^u \frac{ds}{2\pi s}\int_{|z-y|\leq\varepsilon} dz\, \exp\Big\{-\frac{|y|^2}{4s}\Big\}\Big|\frac{|z|^2 - |y|^2}{2s}\Big| \\ &\leq \frac{\varepsilon^3}{2}|y|\int_0^u \frac{ds}{s^2}\, \exp\Big\{-\frac{|y|^2}{4s}\Big\} \\ &\leq \varepsilon^3\, \Psi_1(y),\end{aligned} \tag{2.37}$$

where the function

$$\Psi_1(y) = |y|\int_0^1 \frac{ds}{s^2}\, \exp\Big\{-\frac{|y|^2}{4s}\Big\}$$

is easily seen to be integrable over $\mathbb{R}^2$.

By combining the preceding estimates, we arrive at

$$\mathrm{P}_0\big[\tau_\varepsilon(y) \le u\big] \ge \frac{\pi\varepsilon^2}{f(\varepsilon)} \int_0^u ds\, p_s(y) - \frac{\varepsilon^3}{f(\varepsilon)} \Psi_1(y) \tag{2.38}$$

and using (2.36) it readily follows that

$$\mathrm{P}_0\big[y \in S_\varepsilon(0,1)\big] - \frac{\pi}{|\log\varepsilon|} \int_0^1 ds\, p_s(y) \ge \frac{\varphi_1(y)}{|\log\varepsilon|^2}$$

with a nonnegative function φ_1 such that $\int dy\, \varphi_1(y) < \infty$.

Upper bound. This part is a little more delicate. We rely on the same idea of applying the strong Markov property at time $\tau_\varepsilon(y)$, but we need to be more careful in our estimates. For every $v > 0$, we have

$$\begin{aligned}\mathrm{E}_0\Big[\int_0^{1+v} ds\, \mathbb{I}_{\{|\xi_s - y| \le \varepsilon\}}\Big] &= \mathrm{E}_0\Big[\mathbb{I}_{\{\tau_\varepsilon(y) \le 1+v\}} \mathrm{E}_{\xi_{\tau_\varepsilon(y)}}\Big[\int_0^s dr\, \mathbb{I}_{\{|\xi_r - y| \le \varepsilon\}}\Big]_{s=1+v-\tau_\varepsilon(y)}\Big] \\ &= \mathrm{E}_0\Big[\mathbb{I}_{\{\tau_\varepsilon(y) \le 1+v\}} \int_0^{1+v-\tau_\varepsilon(y)} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big]\Big],\end{aligned}$$

where a_ε is as previously a fixed point of the circle of radius ε centered at the origin. We can rewrite the previous expression as

$$\mathrm{E}_0\Big[\int_0^{1+v} dr\, \mathbb{I}_{\{\tau_\varepsilon(y) \le 1+v-r\}} \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big]\Big] = \int_0^{1+v} dr\, \mathrm{P}_0\big[\tau_\varepsilon(y) \le 1+v-r\big] \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big].$$

We apply this calculation with $v = v_\varepsilon = |\log\varepsilon|^{-1}$. For $r \in [0, v_\varepsilon]$, $\mathrm{P}_0\big[\tau_\varepsilon(y) \le 1 + v_\varepsilon - r\big]$ is bounded from below by $\mathrm{P}_0\big[\tau_\varepsilon(y) \le 1\big]$, and thus

$$\begin{aligned}&\mathrm{P}_0\,\big[\,\tau_\varepsilon(y) \le 1\big] \int_0^{v_\varepsilon} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] \\ &\quad\le \mathrm{E}_0\Big[\int_0^{1+v_\varepsilon} ds\, \mathbb{I}_{\{|\xi_s - y| \le \varepsilon\}}\Big] - \int_{v_\varepsilon}^{1+v_\varepsilon} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] \mathrm{P}_0\big[\tau_\varepsilon(y) \le 1 + v_\varepsilon - r\big].\end{aligned}$$

From the bound (2.37), we have

$$\mathrm{E}_0\Big[\int_0^{1+v_\varepsilon} ds\, \mathbb{I}_{\{|\xi_s - y| \le \varepsilon\}}\Big] \le \pi\varepsilon^2 \int_0^1 ds\, p_s(y) + \varepsilon^3 \Psi_1(y) + v_\varepsilon \varepsilon^2 e^{-|y|^2/10}.$$

On the other hand, by (2.38),

$$\begin{aligned}&\int_{v_\varepsilon}^{1+v_\varepsilon} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] \mathrm{P}_0\big[\tau_\varepsilon(y) \le 1 + v_\varepsilon - r\big] \\ &\quad\ge \int_{v_\varepsilon}^{1+v_\varepsilon} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] \frac{\varepsilon^2}{f(\varepsilon)} \Big(\pi \int_0^{1+v_\varepsilon - r} ds\, p_s(y) - \varepsilon \Psi_1(y)\Big) \\ &\quad= \frac{\pi\varepsilon^2}{f(\varepsilon)} \Big(\int_{v_\varepsilon}^{1+v_\varepsilon} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big]\Big) \Big(\int_0^1 ds\, p_s(y) - \frac{\varepsilon}{\pi} \Psi_1(y)\Big) \\ &\quad\quad - \frac{\pi\varepsilon^2}{f(\varepsilon)} \Big(\int_{v_\varepsilon}^{1+v_\varepsilon} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] \int_{1+v_\varepsilon - r}^1 ds\, p_s(y)\Big).\end{aligned}$$

Straightforward estimates give

$$\int_{v_\varepsilon}^{1+v_\varepsilon} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] = \varepsilon^2 \Big(\frac{1}{2} \log|\log\varepsilon| + \mathcal{O}(1)\Big)$$

and

$$\int_{v_\varepsilon}^{1+v_\varepsilon} dr\ \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] \int_{1+v_\varepsilon - r}^{1} ds\, p_s(y) \le \varepsilon^2 \Psi_2(y),$$

where

$$\Psi_2(y) = \int_0^1 ds\ \log\Big(\frac{1}{1-s}\Big) p_s(y)$$

is integrable over $\mathbb{R}^2$. Summarizing, we have

$$\begin{aligned}
\mathrm{P}_0\big[\tau_\varepsilon(y) \le 1\big] \int_0^{v_\varepsilon} dr\, \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] \\
&\le \Big(\pi\varepsilon^2 \int_0^1 ds\, p_s(y)\Big) \times \bigg(1 - \frac{\big(\frac{1}{2}\log|\log\varepsilon| - K\big)\varepsilon^2}{f(\varepsilon)}\bigg) \\
&\quad + \bigg(\varepsilon^3 + \mathcal{O}\Big(\varepsilon^3 \frac{\log|\log\varepsilon|}{|\log\varepsilon|}\Big)\bigg)\Psi_1(y) + v_\varepsilon \varepsilon^2 e^{-|y|^2/10} + \frac{\pi\varepsilon^4}{f(\varepsilon)}\Psi_2(y).
\end{aligned}$$

Finally, it is easy to verify that

$$g(\varepsilon) := \int_0^{v_\varepsilon} dr\ \mathrm{P}_{a_\varepsilon}\big[|\xi_r| \le \varepsilon\big] \ge \varepsilon^2\big(|\log\varepsilon| - \frac{1}{2}\log|\log\varepsilon| - K'\big),$$

and so by dividing the two sides of the previous inequality by $g(\varepsilon)$, we obtain

$$\mathrm{P}_0\big[\tau_\varepsilon(y) \le 1\big] \le \frac{\pi}{|\log\varepsilon|}\int_0^1 ds\, p_s(y) + \frac{\varphi_2(y)}{|\log\varepsilon|^2},$$

with a function φ_2 such that $\int \varphi_2(y)dy < \infty$. This completes the proof of Lemma 2.13. □

CHAPTER 3

Escape probabilities for branching Brownian motion among mild obstacles

We investigate the behaviour of critical branching Brownian motion killed at rate $\varepsilon \in (0,1)$ within a random collection of obstacles distributed according to a Poisson point process on $\mathbb{R}^d$ ($d \geq 1$). Fixing a domain $A \subset \mathbb{R}^d$ containing 0, we focus on the probability that our branching Brownian motion, starting from one particle at the origin, hits $(RA)^c$ for a constant $R \gg 1$. We prove that, almost surely for the law of the obstacles, this probability behaves like a constant times R^{-2} as R tends to infinity and ε tends to zero. This constant depends on the limit of εR^2 and can be identified in terms of the solution of a semilinear partial differential equation with singular boundary conditions. The proof uses a quenched convergence result showing that as ε tends to zero, a suitable rescaling of the branching Brownian motion killed at rate ε within the obstacles converges in distribution to super-Brownian motion killed at a finite rate homogeneously over $\mathbb{R}^d$. When $\varepsilon R^2 \to \infty$, the constant of interest is zero and we complement our analysis by proving the exponential decay of the corresponding hitting probabilities. The results of this chapter are a joint work with Jean-François Le Gall.

Contents

1. Introduction

In the present work, we are interested in the long-term behaviour of branching Brownian motion killed in Poissonian obstacles. Let us start by describing a simple special case of our results. We consider a critical branching Brownian motion in $\mathbb{R}^d$ ($d \geq 1$), where all initial particles start from the origin. We assume that particles are killed at a (small) rate $\varepsilon > 0$ within random balls of fixed radius, whose centres are distributed according to a homogeneous Poisson point process on $\mathbb{R}^d$. Then, how many initial particles do we need so that, with high probability, one of their descendants reaches distance R from the origin ? Let $p_\varepsilon(R)$ be the (quenched) probability for our randomly killed branching Brownian motion starting with a single particle at 0 to visit the complement of a large ball of radius R centered at the origin. The preceding question is equivalent to determining the limiting behaviour of $p_\varepsilon(R)$ when ε tends to 0 and simultaneously R tends to infinity.

The answer involves several regimes depending on the respective values of ε and R. If ε is small in comparison with $1/R^2$, the killing phenomenon does not matter and the result is the same as if there were no killing: $p_\varepsilon(R)$ behaves like a constant times $1/R^2$ (informally, the branching process must survive up to a time of order R^2 so that at least one of the particles travels a distance R, and well-known estimates for critical branching processes then lead to the correct asymptotics). On the other hand, if ε is large in comparison with $1/R^2$, then the probability $p_\varepsilon(R)$ decreases exponentially fast as a function of $R\sqrt{\varepsilon}$: See Proposition 3.2 below.

Our main results focus on the critical regime where εR^2 converges to a constant $a > 0$. We show that the probability $p_\varepsilon(R)$ behaves like R^{-2}, as in the case without killing, but with a multiplicative constant which depends on a and can be identified as the value at the origin of the solution of a semilinear partial differential equation with singular boundary conditions. A key tool to derive these asymptotics is a quenched homogenization theorem which shows that our branching Brownian motions among obstacles, suitably rescaled, are close to super-Brownian motion killed at a certain rate depending on a.

Let us formulate our assumptions more precisely in order to state our results. First, let us define the collection of obstacles. We denote the set of all compact subsets of $\mathbb{R}^d$ by $\mathcal{K}$. This set is equipped with the usual Hausdorff metric d_H. Recall that $(\mathcal{K}, d_H)$ is a Polish space. For every $r > 0$, $\mathcal{K}_r$ denotes the subset of $\mathcal{K}$ which consists of all compact sets that are contained in the closed ball of radius r centered at the origin. Let Θ be a finite measure on $\mathcal{K}$, and assume that Θ is supported on $\mathcal{K}_{r_0}$ for some $r_0 > 0$. Let

$$\mathcal{N} = \sum_{i\in I} \delta_{(x_i,K_i)}$$

be a Poisson point measure on $\mathbb{R}^d \times \mathcal{K}$ with intensity $\lambda_d \otimes \Theta$, where λ_d stands for Lebesgue measure on $\mathbb{R}^d$. We assume that this point measure is defined on a probability space $(\mathbf{\Omega}, \mathbf{P})$ and we denote the generic element of $\mathbf{\Omega}$ by ϖ. Our set of obstacles is then defined by

$$\Gamma_\varpi = \bigcup_{i\in I} (x_i + K_i),$$

where obviously $x_i + K_i = \{z = x_i + y : y \in K_i\}$. Note that we use the notation Γ_ϖ to emphasize that the set of obstacles depends on the variable ϖ representing the environment. Let us also define a constant κ by

$$\kappa = \mathbf{P}(0 \in \Gamma_\varpi) = 1 - \exp\Big(- \int_{\mathcal{K}} \Theta(dK)\, \lambda_d(K)\Big).$$

To avoid trivial cases, we assume that $\kappa > 0$, or equivalently $\Theta(\lambda_d(K) > 0) > 0$. By translation invariance, we also have $\mathbf{P}(x \in \Gamma_\varpi) = \kappa$ for every $x \in \mathbb{R}^d$.

Let us now introduce the sequence of branching Brownian motions of interest. Given $\varpi \in \mathbf{\Omega}$ and a parameter $\varepsilon \geq 0$, we consider a branching Brownian motion on $\mathbb{R}^d$ such that

- each particle moves around in $\mathbb{R}^d$ according to the law of Brownian motion killed at rate ε within Γ_ϖ ;
- each particle branches at rate 1. During a branching event, the particle generates a random number of offspring, according to an offspring distribution ν which has mean one and finite variance $\sigma^2 \in (0, \infty)$.

This branching Brownian motion is denoted by $Z^{\varpi,\varepsilon} = (Z_t^{\varpi,\varepsilon})_{t\geq 0}$, where $Z_t^{\varpi,\varepsilon}$ stands for the sum of the Dirac point masses at the particles alive at time t. The processes $Z^{\varpi,\varepsilon}$ are defined on a probability space Ω. For every finite point measure μ on $\mathbb{R}^d$, we use the notation $\mathbb{P}_\mu$ for the probability measure on Ω under which each of the processes $Z^{\varpi,\varepsilon}$ starts from μ.

Let A be a bounded domain of class C^2 in $\mathbb{R}^d$ containing 0. We say that the branching Brownian motion $Z^{\varpi,\varepsilon}$ hits A^c if there exists $t > 0$ such that $Z_t^{\varpi,\varepsilon}(A^c) > 0$. We are interested in asymptotics for the quantity

$$\mathbb{P}_{\delta_0}(Z^{\varpi,\varepsilon} \text{ hits } (RA)^c)$$

when $R \to \infty$ and $\varepsilon \to 0$. Here we use the obvious notation $RA = \{z = Ry : y \in A\}$.

THEOREM 3.1. *For every $a \geq 0$, let $u_{(a)} = (u_{(a)}(x), x \in A)$ be the unique nonnegative solution of the singular boundary value problem*

$$\begin{cases} \dfrac{1}{2}\Delta u = \dfrac{\sigma^2}{2} u^2 + a\, u & \text{in } A, \\ u_{|\partial A} = +\infty\ . \end{cases} \tag{3.1}$$

Then,

$$\lim_{R\to\infty} \Big(\sup_{\varepsilon \geq 0} \Big| R^2 \mathbb{P}_{\delta_0}(Z^{\varpi,\varepsilon} \text{ hits } (RA)^c) - u_{(\kappa\varepsilon R^2)}(0) \Big| \Big) = 0\ , \qquad \mathbf{P}(d\varpi) \text{ a.s.}$$

It is not hard to see that $u_{(a)}(0)$ tends to 0 as $a \to \infty$ (see Lemma 3.7 below), and thus Theorem 3.1 does not give much information when $R \to \infty$ and $\varepsilon \to 0$ in such a way that εR^2 tends to ∞. In that case, the next proposition provides an exponential decay, which contrasts with the preceding theorem. Since our bounds are clearly not optimal, we consider only the case when A is a ball. We denote the open ball of radius r centered at the origin by $B(0, r)$. In the general case we may apply the bounds (i) and (ii) of the proposition after replacing A by a ball $B(0, r)$ such that $B(0, r) \supset A$ or $B(0, r) \subset A$ respectively.

PROPOSITION 3.2. (i) *There exists a positive constant $C_0 = C_0(\nu)$ such that, for every $\varpi \in \mathbf{\Omega}$, $R \geq 1$ and $\varepsilon \in [1/R^2, 1]$,*

$$\mathbb{P}_{\delta_0}(Z^{\varpi,\varepsilon} \text{ hits } B(0,R)^c) \geq C_0 \varepsilon \exp(-R\sqrt{2\varepsilon}).$$

(ii) *There exists a positive constant $C_1 = C_1(\nu, \Theta)$ such that for every $R \geq 1$ and $\varepsilon \in [1/R^2, 1]$,*

$$\mathbf{P} \otimes \mathbb{P}_{\delta_0}(Z^{\varpi,\varepsilon} \text{ hits } B(0,R)^c) \leq \exp(-C_1 R\sqrt{\varepsilon}).$$

Part (i) of the proposition is derived from an estimate about branching Brownian motion killed homogeneously at rate ε, which explains why this bound holds for every $\varpi \in \mathbf{\Omega}$ and does not depend on the measure Θ.

Proposition 3.2 only gives rather crude estimates, and it would be of interest to obtain more precise information on the decay of the quenched probabilities $\mathbb{P}_{\delta_0}(Z^{\varpi,\varepsilon} \text{ hits } (RA)^c)$ in the case when εR^2 tends to ∞. This leads to large deviations problems in the spirit of the work of Sznitman [**Szn98**], which we do not address here.

A major ingredient of the proof of Theorem 3.1 is the following quenched homogenization result. We need to introduce a rescaled version of the process $Z^{\varpi,\varepsilon}$. For every $\varepsilon > 0$ and every $t \geq 0$, let us define a random measure $X_t^{\varpi,\varepsilon}$ on $\mathbb{R}^d$ by setting

$$\int X_t^{\varpi,\varepsilon}(dx)\, \varphi(x) = \varepsilon \int Z^{\varpi,\varepsilon}_{\varepsilon^{-1}t}(dx)\, \varphi(\varepsilon^{1/2}x).$$

for every bounded measurable function φ on $\mathbb{R}^d$.

For every real $x \geq 0$, $[x]$ denotes the integer part of x.

THEOREM 3.3. *Except for a $\mathbf{P}$-negligible set of values of ϖ, the law of $(X_t^{\varpi,\varepsilon})_{t\geq 0}$ under $\mathbb{P}_{[\varepsilon^{-1}]\delta_0}$ converges weakly as $\varepsilon \to 0$, in the Skorokhod sense, to that of a super-Brownian motion with branching mechanism $\psi_{(\kappa)}(u) := \frac{\sigma^2}{2}u^2 + \kappa u$ started at δ_0.*

The definition of super-Brownian motion with branching mechanism $\psi_{(\kappa)}$ is recalled in Section 2 below.

As a hint of why Theorem 3.3 should be true, notice that for a given realization of the obstacles, the probability that a single Brownian motion starting from 0 and killed at rate ε within Γ_ϖ is still alive by time $t > 0$ is given by

$$\mathrm{E}\Big[\exp\Big\{-\varepsilon\int_0^t \mathbb{I}_{\Gamma_\varpi}(\xi_s)ds\Big\}\Big], \tag{3.2}$$

where ξ denotes standard d-dimensional Brownian motion. Let us focus on the integral within the exponential in (3.2). Averaging over the law of the obstacles and using Fubini's theorem, we obtain for each $t \geq 0$

$$\mathbf{E}\Big[\mathrm{E}\Big[\varepsilon\int_0^t \mathbb{I}_{\Gamma_\varpi}(\xi_s)ds\Big]\Big] = \varepsilon\,\mathrm{E}\Big[\int_0^t \mathbf{P}[\xi_s\in\Gamma_\varpi]ds\Big] = \varepsilon\kappa t.$$

This observation suggests that when ε is small, the killing will be felt by our Brownian motion only after a time of order ε^{-1}, by which a particle has traveled a distance of the order of $\varepsilon^{-1/2}$. This gives us the suitable rescaling to consider. Besides, the quantity in (3.2) written at time $\varepsilon^{-1}t$ should be close to $e^{-\kappa t}$. We can thus guess, and easily prove, that the rescaled Brownian motion $(\sqrt{\varepsilon}\,\xi_{\varepsilon^{-1}t},\ t\geq 0)$, killed at rate 1 within $\sqrt{\varepsilon}\,\Gamma_\varpi$, converges to Brownian motion killed at homogeneous rate κ as $\varepsilon \to 0$. Theorem 3.3 shows that an analogous convergence indeed holds in our more general framework of branching Brownian motions, for any fixed ϖ contained in a set of $\mathbf{P}$-probability one.

Let us briefly explain how Theorem 3.1 is derived from Theorem 3.3. Consider a sequence (ε_n, R_n) such that $R_n \to \infty$ and $\varepsilon_n R_n^2$ converges to a positive constant a. By a simple scaling transformation, the probability $\mathbb{P}_{\delta_0}(Z^{\varpi,\varepsilon_n}$ hits $(R_nA)^c)$ coincides with $\mathbb{P}_{\delta_0}(X^{\varpi,\varepsilon_n}$ hits $(b_nA)^c)$, where $b_n = \varepsilon_n^{1/2}R_n$ converges to $\sqrt{a}$. We can then use Theorem 3.3 to investigate the asymptotic behaviour of the latter hitting probabilities. This limiting behaviour involves the corresponding hitting probabilities for super-Brownian motion, which are known to be related to solutions of semilinear partial differential equations from the work of Dynkin [**Dy91a, Dy93**]. One difficulty in implementing the preceding idea comes from the fact that the convergence in Theorem 3.3 is not strong enough to ensure that hitting probabilities for the processes $(X_t^{\varpi,\varepsilon})_{t\geq 0}$ converge to hitting probabilities for the limiting process. Much of the proof of Theorem 3.1 in Section 4 is devoted to a precise justification of this property (Lemma 3.13).

To complete this introduction, let us mention that branching Brownian motion and superprocesses among random obstacles have been studied recently in several papers, including Engländer and den Hollander [**EdH03**] and Engländer [**Eng08**]. These papers concentrate on the case of supercritical branching, in contrast with critical branching which is considered here. See also the survey [**Eng07**]. A homogenization theorem related to Theorem 3.3 has been proved in [**Véb09**] for super-Brownian motion among hard obstacles, in the case when the intensity of the obstacles grows to infinity but their diameters shrink to 0. There is a huge literature about Brownian motion and random walks among (hard or mild) obstacles, and the reader may look at the book of Sznitman [**Szn98**] for additional references.

The rest of this chapter is laid out as follows. In Section 2, we introduce the basic notation and objects, and state several results about hitting probabilities for spatial branching processes we shall need in the sequel. Theorem 3.3 and Proposition 3.2 are proved in Section 3. Theorem 3.1 is then derived in Section 4.

2. Preliminaries

2.1. Notation. We denote the set of all finite measures on $\mathbb{R}^d$ by $\mathcal{M}_f(\mathbb{R}^d)$. This set is equipped with the weak topology. We write $\mathcal{M}_p(\mathbb{R}^d)$ for the subset of $\mathcal{M}_f(\mathbb{R}^d)$ which consists of all finite point measures on $\mathbb{R}^d$. If $\mu \in \mathcal{M}_f(\mathbb{R}^d)$ and f is a function on $\mathbb{R}^d$, we write $\langle \mu, f\rangle$ for the integral $\int f\, d\mu$ whenever this integral makes sense.

If E is a metric space, we denote the set of all bounded continuous functions on the space E by $\bar{C}(E)$ and we let $\|f\|$ stand for the supremum norm of $f \in \bar{C}(E)$. We write $C^2(\mathbb{R}^d)$ for the set of all twice continuously differentiable functions on $\mathbb{R}^d$, and $\bar{C}^2(\mathbb{R}^d)$ for that of all bounded functions in $C^2(\mathbb{R}^d)$ whose first and second derivatives are also bounded. An index $+$ added to this notation means that we require the functions to be nonnegative. We equip $\bar{C}^2(\mathbb{R}^d)$ with the topology induced by the seminorms $\|f\|_{(R)}$, where for every $R > 0$

$$\|f\|_{(R)} := \sup_{|x|\le R} \Big\{ |f(x)| + \sum_{i=1}^d \Big|\frac{\partial f}{\partial x_i}(x)\Big| + \sum_{i,j=1}^d \Big|\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\Big|\Big\}.$$

If E is a Polish space, we let $D_E[0,\infty)$ be the set of all càdlàg paths with values in E, equipped with the Skorokhod topology.

If $x \in \mathbb{R}^d$ and $r > 0$, $B(x,r)$ denotes the open ball of radius r centered at x, and $\bar{B}(x,r)$ stands for the corresponding closed ball. More generally, the closure of a subset F of $\mathbb{R}^d$ is denoted by $\bar{F}$. Lebesgue measure on $\mathbb{R}^d$ is denoted by λ_d.

Finally, the notation $\xi = (\xi_t)_{t\ge 0}$ will stand for a standard Brownian motion in $\mathbb{R}^d$, which starts from x under the probability measure P_x. It will also be convenient to use the notation $\xi^{\varpi,\varepsilon}$ for Brownian motion in $\mathbb{R}^d$ killed at rate ε in the set Γ_ϖ. As usual, the value of $\xi^{\varpi,\varepsilon}$ after its killing time is a cemetery point Δ added to $\mathbb{R}^d$, and we agree that all functions vanish at Δ.

2.2. Super-Brownian motion. Let $a \ge 0$ and set $\psi_{(a)}(u) = \frac{\sigma^2}{2}u^2 + a\,u$, for every $u \ge 0$ (the offspring distribution ν, and thus the parameter $\sigma > 0$ are fixed throughout this work). Super-Brownian motion with branching mechanism $\psi_{(a)}$ is the continuous strong Markov process with values in $\mathcal{M}_f(\mathbb{R}^d)$, whose transition kernels $(Q_t)_{t\ge 0}$ are characterized as follows: For every $g \in \bar{C}_+(\mathbb{R}^d)$ and every $\mu \in \mathcal{M}_f(\mathbb{R}^d)$, we have for every $t \ge 0$

$$\int Q_t(\mu, d\mu') \exp(-\langle \mu', g\rangle) = \exp(-\langle \mu, V_t g\rangle), \tag{3.3}$$

where the function $u_t(x) = V_t g(x)$, $t \ge 0$, $x \in \mathbb{R}^d$, is the unique nonnegative solution of the semilinear parabolic problem

$$\begin{cases} \dfrac{\partial u}{\partial t} = \frac{1}{2}\Delta u - \psi_{(a)}(u) & \text{in } (0,\infty)\times\mathbb{R}^d\,, \\ u_0 = g\,. \end{cases}$$

Let $Y = (Y_t)_{t\ge 0}$ be a super-Brownian motion with branching mechanism $\psi_{(a)}$, started at $\mu \in \mathcal{M}_f(\mathbb{R}^d)$. Then, for every $g \in \bar{C}^2_+(\mathbb{R}^d)$

$$e^{-\langle Y_t, g\rangle} - e^{-\langle Y_0, g\rangle} - \int_0^t \Big\langle Y_s, -\frac{1}{2}\Delta g + \psi_{(a)}(g)\Big\rangle\, e^{-\langle Y_s, g\rangle} ds \tag{3.4}$$

is a martingale. It is well known that this martingale problem and the initial value μ characterize the law of Y. This is indeed an application of the classical "duality method" (see in particular Chapter 4 in [**EK86**]). The nonlinear semigroup $g \to V_t g$ provides a deterministic dual to super-Brownian motion, and the duality argument then shows that if a measure-valued process started from μ satisfies the preceding martingale problem, the Laplace functional of its value at time t must be given by the right-hand side of (3.3). See Section 1.6 of [**Eth00**] for more details.

2.3. Branching Brownian motion among random obstacles. In view of our applications (and in particular because we want to refer to some results of [**Ch91**]), it will be convenient to give a more formal description of the branching Brownian motions that were already introduced in Section 1 above. Recall that our offspring distribution ν is assumed to be critical and that $\mathrm{var}(\nu) = \sigma^2 \in (0, \infty)$. The probability generating function of ν will be denoted by Υ.

Let $\mathcal{T}$ be a Galton-Watson tree with offspring distribution ν (see e.g. [**LG05**]). As usual, we view $\mathcal{T}$ as a random finite subset of

$$U := \bigcup_{n=0}^{\infty} \mathbb{N}^n,$$

where $\mathbb{N}^0 = \{\varnothing\}$. If $v = (v_1, \ldots, v_n) \in U \backslash \{\varnothing\}$, the parent of v is denoted by $\hat{v} = (v_1, \ldots, v_{n-1})$ and we also use the notation $v \prec v'$ to mean that v' is a descendant of v distinct from v. Consider a collection $(e_v, v \in U)$ of independent exponential random variables with parameter 1, which is also independent of $\mathcal{T}$. We define for every $v \in U$ its birth time α_v and its death time β_v recursively by setting $\alpha_\varnothing = 0$ and $\beta_\varnothing = e_\varnothing$, and for every $v \in U \backslash \{\varnothing\}$,

$$\alpha_v = \beta_{\hat{v}} \ , \ \beta_v = \alpha_v + e_v.$$

Let us now construct the spatial motions. Fix a starting point $x \in \mathbb{R}^d$, and consider a collection $(B^v, v \in U)$ of independent standard Brownian motions in $\mathbb{R}^d$ (started from 0), independent of $\mathcal{T}$ and of $(e_v, v \in U)$. For every $v \in U$, define the *historical path* $\omega^v = (\omega^v_t, 0 \le t \le \beta_v)$ associated with v in the following way. First $\omega^\varnothing_t = x + B^\varnothing_t$ for $0 \le t \le \beta_\varnothing$. Then, if $v \in U \backslash \{\varnothing\}$, set $\omega^v_t = \omega^{\hat{v}}_t$ for all $0 \le t \le \alpha_v$ and

$$\omega^v_t = \omega^{\hat{v}}_{\alpha_v} + B^v_{t-\alpha_v} \qquad \text{for } \alpha_v \le t \le \beta_v.$$

A branching Brownian motion (without killing in obstacles) starting from δ_x is obtained by setting for every $t \ge 0$,

$$Z_t = \sum_{v \in \mathcal{T}, v \sim t} \delta_{\omega^v_t},$$

where the notation $v \sim t$ means that $\alpha_v \le t < \beta_v$.

In this formalism, it is now easy to introduce killing in obstacles. Consider yet another independent collection $(\gamma_v)_{v \in U}$ of independent exponential random variables with parameter 1, and define for every $\varpi \in \mathbf{\Omega}$, $\varepsilon \ge 0$, and for every $v \in U$

$$\zeta^{\varpi,\varepsilon}_v := \inf\left\{ s \in [\alpha_v, \beta_v) : \int_{\alpha_v}^{s} dr\, \mathbb{I}_{\Gamma_\varpi}(\omega^v_r) > \varepsilon^{-1} \gamma_v \right\},$$

where $\inf \varnothing = \infty$. By setting

$$Z^{\varpi,\varepsilon}_t = \sum_{v \in \mathcal{T}, v \sim t} \mathbb{I}_{\{t < \zeta^{\varpi,\varepsilon}_v \text{ and } \zeta^{\varpi,\varepsilon}_{v'} = \infty, \text{ for every } v' \prec v\}} \delta_{\omega^v_t},$$

we obtain a branching Brownian motion killed at rate ε in the obstacle set Γ_ϖ, starting from δ_x. An obvious extension of the preceding construction allows us to obtain branching Brownian motions starting from any point measure $\mu \in \mathcal{M}_p(\mathbb{R}^d)$.

We now recall a special case of the classical convergence of rescaled branching Brownian motions towards super-Brownian motion. For our applications, we state the case where particles are killed at a constant rate homogeneously over $\mathbb{R}^d$ (this case is obtained from the preceding construction of $Z^{\varpi,\varepsilon}$ by replacing Γ_ϖ by $\mathbb{R}^d$). In the next two statements, for every $\varepsilon > 0$, $Z^{(\varepsilon)} = (Z^{(\varepsilon)}_t)_{t \ge 0}$ denotes a branching Brownian motion with offspring distribution ν, where particles are killed homogeneously over $\mathbb{R}^d$ at rate ε. As previously $Z^{(\varepsilon)}$ starts from μ under the probability measure $\mathbb{P}_\mu$, for every $\mu \in \mathcal{M}_p(\mathbb{R}^d)$.

PROPOSITION 3.4. *Let $a \geq 0$. For every $\varepsilon > 0$, define a measure-valued process $(X_t^{(\varepsilon)})_{t\geq 0}$ by setting*

$$\langle X_t^{(\varepsilon)}, \varphi\rangle = \varepsilon \int Z^{(a\varepsilon)}_{\varepsilon^{-1}t}(dx)\,\varphi(\varepsilon^{1/2}x),$$

for every $\varphi \in \bar{C}(\mathbb{R}^d)$. For every fixed $\eta > 0$, the law of $X^{(\varepsilon)}$ under $\mathbb{P}_{[\eta\varepsilon^{-1}]\delta_0}$ converges as $\varepsilon \to 0$, in the Skorokhod sense, towards the law of super-Brownian motion with branching mechanism $\psi_{(a)}$ starting from $\eta\delta_0$.

A proof of Proposition 3.4 can be found in Chapter 1 of [**Eth00**] in the case $a = 0$, and arguments are easily adapted to cover the general case.

Finally, we shall use an estimate for the probability that a branching Brownian motion starting from δ_0 exits a large ball centered at the origin. Similar estimates can be found in Sawyer and Fleischman [**SF79**], but we provide a short proof for the sake of completeness.

LEMMA 3.5. *Suppose that $d = 1$. There exist two positive constants $C_0' = C_0'(\nu)$ and $C_1' = C_1'(\nu)$ such that, for every $\varepsilon \in [0,1]$ and $r \geq 1$,*

$$C_0'\Big(r^{-2}\,\mathbb{I}_{\{r\leq \frac{1}{\sqrt{\varepsilon}}\}} + \varepsilon e^{-r\sqrt{2\varepsilon}}\,\mathbb{I}_{\{r>\frac{1}{\sqrt{\varepsilon}}\}}\Big) \leq \mathbb{P}_{\delta_0}(Z^{(\varepsilon)} \text{ hits } (-r,r)^c)$$
$$\leq C_1'\Big(r^{-2}\,\mathbb{I}_{\{r\leq \frac{1}{\sqrt{\varepsilon}}\}} + \varepsilon e^{-r\sqrt{2\varepsilon}}\,\mathbb{I}_{\{r>\frac{1}{\sqrt{\varepsilon}}\}}\Big).$$

REMARK 3.6. *As an immediate consequence of the upper bound of the lemma, we have in dimension d, for every $r > 0$,*

$$\mathbb{P}_{\delta_0}(Z^{(0)} \text{ hits } B(0,r)^c) \leq C_1''\,(r+1)^{-2} \tag{3.5}$$

with a constant $C_1'' = C_1''(d,\nu)$.

Proof. It clearly suffices to prove that the stated bounds hold for the quantity $\mathbb{P}_{\delta_r}(Z^{(\varepsilon)}$ hits $(-\infty,0])$ instead of $\mathbb{P}_{\delta_0}(Z^{(\varepsilon)}$ hits $(-r,r)^c)$. We fix $\varepsilon \geq 0$, and for every $x > 0$ and $t \geq 0$, we set

$$q_\varepsilon(x,t) = \mathbb{P}_{\delta_x}\left(Z^{(\varepsilon)} \text{ does not hit } (-\infty,0] \text{ before time } t\right).$$

and

$$p_\varepsilon(x) = \mathbb{P}_{\delta_x}\left(Z^{(\varepsilon)} \text{ hits } (-\infty,0]\right) = \lim_{t\uparrow\infty} \uparrow (1 - q(x,t)).$$

In this proof only, we write P_x^ε for the probability under which ξ is a Brownian motion starting from x and killed at rate ε (upon killing, ξ is sent to the cemetery point Δ and as usual we agree that all functions vanish at Δ). Write $S := \inf\{t \geq 0 : \xi_t \in (-\infty,0]\}$. By standard arguments (see e.g. the proof of Proposition II.3 in [**LG99**]), the function q_ε solves the integral equation

$$q_\varepsilon(x,t) = \mathrm{P}_x^\varepsilon(S > t) + \mathrm{E}_x^\varepsilon\Big[\int_0^{t\wedge S} (\Upsilon(q_\varepsilon(\xi_s,t-s)) - q_\varepsilon(\xi_s,t-s))\,ds\Big],$$

where we recall that Υ denotes the generating function of the offspring distribution ν. For every $a \in [0,1]$, set $\Phi(a) = \Upsilon(1-a) - (1-a)$. Note that $\Phi(0) = 0$ and the function Φ is monotone increasing under our assumptions. Furthermore, $\Phi(a) = \frac{\sigma^2}{2}a^2 + o(a^2)$ when $a \to 0$. By a monotone passage to the limit we get that, for every $x > 0$,

$$p_\varepsilon(x) + \mathrm{E}_x^\varepsilon\Big[\int_0^S \Phi(p_\varepsilon(\xi_s))\,ds\Big] = \mathrm{P}_x^\varepsilon(S < \infty). \tag{3.6}$$

It follows that the function p_ε satisfies the differential equation

$$\frac{1}{2}p_\varepsilon'' = \varepsilon\, p_\varepsilon + \Phi(p_\varepsilon)$$

on $(0,\infty)$ with boundary conditions $p_\varepsilon(0) = 1, p_\varepsilon(\infty) = 0$. By solving this differential equation, we get, for every $x > 0$,

$$\int_{p_\varepsilon(x)}^{1} \frac{du}{\sqrt{2\varepsilon u^2 + 4\Gamma(u)}} = x$$

where $\Gamma(u) = \int_0^u \Phi(v)\,dv$. Note that there exist positive constants c, c' such that $cu^3 \leq \Gamma(u) \leq c'u^3$ for every $u \in [0,1]$. The desired bounds then follow from easy analytic arguments. $\square$

2.4. Hitting probabilities for super-Brownian motion. Let $Y^{(a)} = (Y_t^{(a)})_{t\geq 0}$ be a super-Brownian motion with branching mechanism $\psi_{(a)}$ for some $a \geq 0$. Suppose that $Y^{(a)}$ starts from μ under the probability measure P_μ, for every $\mu \in \mathcal{M}_f(\mathbb{R}^d)$.

The range of $Y^{(a)}$ is by definition

$$\mathcal{R}(Y^{(a)}) = \bigcup_{\varepsilon>0} \left(\overline{\bigcup_{t=\varepsilon}^{\infty} \operatorname{supp}(Y_t^{(a)})} \right),$$

where for every $\mu \in \mathcal{M}_f(\mathbb{R}^d)$, $\operatorname{supp}(\mu)$ denotes the topological support of μ.

Let D be a domain in $\mathbb{R}^d$ and let $x \in D$. Consider the process $Y^{(a)}$ started from δ_x. We say that $Y^{(a)}$ hits D^c if the range $\mathcal{R}(Y^{(a)})$ intersects D^c. By a famous result of Dynkin [**Dy91a, Dy93**] the function

$$u_{(a)}^D(x) = -\log\Big(1 - P_{\delta_x}(\mathcal{R}(Y^{(a)}) \cap D^c \neq \emptyset)\Big)\,, \quad x \in D\,,$$

is the maximal nonnegative solution of the semilinear partial differential equation $\frac{1}{2}\Delta u = \psi_{(a)}(u)$ in D.

Under mild regularity assumptions on D (which hold e.g. when D satisfies an exterior cone condition at every point of ∂D), the function $u_{(a)}^D$ has boundary value $+\infty$ at every point of ∂D and is the unique nonnegative solution of the equation $\frac{1}{2}\Delta u = \psi_{(a)}(u)$ in D with boundary value $+\infty$ everywhere on ∂D. A discussion of this result and related ones can be found in Chapter VI of the book [**LG99**]. This reference considers only the case $a = 0$, but the same results can be obtained for any $a \geq 0$ by similar arguments: Note that the Brownian snake approach can be extended from the case $a = 0$ considered in [**LG99**] to $a \geq 0$, simply by replacing the reflecting Brownian motion driving the snake by a reflecting Brownian motion with negative drift (see Chapter 4 of [**DLG02**] for a discussion of the snake approach to superprocesses with a general branching mechanism).

We shall be interested in the special case $D = A$. Recall that $0 \in A$ and that we assume A is a domain of class C^2, meaning that the boundary of A can be represented locally as the graph of a twice continuously differentiable function, in a suitable system of coordinates. We write $u_{(a)}(x) = u_{(a)}^A(x)$ to simplify notation. From the analytic viewpoint, the function $u_{(a)}$ may be constructed as follows. For every integer $n \geq 1$, let $u_{(a),n}$ be the unique nonnegative solution of the nonlinear Dirichlet problem

$$\begin{cases} \frac{1}{2}\Delta u = \psi_{(a)}(u) & \text{in } A\,, \\ u_{|\partial A} = n\,. \end{cases}$$

Then $u_{(a)} = \lim \uparrow u_{(a),n}$ as $n \to \infty$.

The following lemma records certain analytic properties which will be useful in the forthcoming proofs.

LEMMA 3.7. (i) *Let* $x \in A$. *The function* $a \mapsto u_{(a)}(x)$ *is continuous on* $[0, \infty)$, *and tends to* 0 *as* $a \to \infty$.
(ii) *For every* $\delta \in (0, \mathrm{dist}(0, A^c))$, *let* A_δ *be the subdomain of* A *defined as the connected component of the open set* $\{x \in A : \mathrm{dist}(x, A^c) > \delta\}$ *that contains* 0. *Then, for every* $a \geq 0$, $u^{A_\delta}_{(a)}(0)$ *tends to* $u_{(a)}(0)$ *as* $\delta \to 0$.

Proof. (i) We first observe that the function $a \mapsto u_{(a)}(x)$ is monotone nonincreasing, for every $x \in A$. To see this, apply a standard comparison principle (see e.g. Lemma V.7 in [**LG99**]) to obtain that $u_{(a'),n} \leq u_{(a),n}$ if $a \leq a'$, for every $n \geq 1$, and let $n \to \infty$.

Let then $(a_k)_{k\geq 1}$ be a sequence of nonnegative reals increasing to $a \in (0, \infty)$. We can set for every $x \in A$

$$v(x) = \lim_{k\uparrow\infty} \downarrow u_{(a_k)}(x),$$

and we have $v \geq u_{(a)}$. In order to verify that $v \leq u_{(a)}$, we only need to check that v solves $\frac{1}{2}\Delta v = \psi_{(a)}(v)$ in A (recall that $u_{(a)}$ is the maximal nonnegative solution of this equation). To do so, let B be an open ball whose closure $\bar{B}$ is contained in A. For every $k \geq 1$, the restriction of $u_{(a_k)}$ to B solves the equation $\frac{1}{2}\Delta u = \psi_{(a_k)}(u)$ in B. By the probabilistic interpretation of the integral equation associated with this PDE (see e.g. Chapter V in [**LG99**]), this implies that, for every $x \in B$,

$$u_{(a_k)}(x) + \mathrm{E}_x\left[\int_0^{\tau_B} \psi_{(a_k)}(u_{(a_k)}(\xi_s))\, ds\right] = \mathrm{E}_x\big[u_{(a_k)}(\xi_{\tau_B})\big],$$

where we recall our notation ξ for a Brownian motion starting from x under the probability measure P_x, and $\tau_B := \inf\{t \geq 0 : \xi_t \notin B\}$. By passing to the limit $k \to \infty$ in the previous display, we can write

$$v(x) + \mathrm{E}_x\left[\int_0^{\tau_B} \psi_{(a)}(v(\xi_s))\, ds\right] = \mathrm{E}_x\big[v(\xi_{\tau_B})\big],$$

which is enough to obtain that v solves $\frac{1}{2}\Delta v = \psi_{(a)}(v)$ in B, and therefore in A since B was arbitrary.

Similar arguments show that, if $(a_k)_{k\geq 1}$ is a decreasing sequence of nonnegative reals converging to $a \in [0, \infty)$, then $u_{(a_k)}(x)$ converges to $u_{(a)}(x)$ for every $x \in A$. Finally, the fact that $u_{(a)}(x)$ tends to 0 as $a \to \infty$ can be obtained from the comparison principle: If B is a ball such that $\bar{B} \subset A$, the restriction of $u_{(a)}$ to B is bounded above by the solution $v_{(a)}$ of the linear equation $\frac{1}{2}\Delta v_{(a)} = a\, v_{(a)}$ in B, with boundary value equal to the restriction of $u_{(0)}$ to B. It is easily seen that $v_{(a)}(x) \mapsto 0$ as $a \to \infty$, for instance by using the Feynman-Kac formula.
(ii) Fix $a \geq 0$. If $0 < \delta < \delta'$, the closure of $A_{\delta'}$ is contained in A_δ. The restriction of $u^{A_\delta}_{(a)}$ to $A_{\delta'}$ is a nonnegative solution of $\frac{1}{2}\Delta u = \psi_{(a)}(u)$ in $A_{\delta'}$ and is thus bounded above by $u^{A_{\delta'}}_{(a)}$. Hence, for every fixed $x \in A$ the function $\delta \to u^{A_\delta}_{(a)}(x)$, which is defined for $\delta > 0$ small enough, is nondecreasing and we can set

$$v(x) = \lim_{\delta\downarrow 0} u^{A_\delta}_{(a)}(x).$$

By the same argument we used to obtain the monotonicity of the mapping $\delta \to u^{A_\delta}_{(a)}(x)$, we also have $v(x) \geq u_{(a)}(x)$ for every $x \in A$. To obtain the reverse inequality $v \leq u_{(a)}$, it is enough to verify that v solves $\frac{1}{2}\Delta v = \psi_{(a)}(v)$ in B. But this follows by arguments similar to those we used in the proof of part (i) of the lemma. □

LEMMA 3.8. *For every* $a \geq 0$ *and* $x \in A$,

$$\{\mathcal{R}(Y^{(a)}) \cap A^c \neq \emptyset\} = \{\mathcal{R}(Y^{(a)}) \cap \bar{A}^c \neq \emptyset\}\,, \quad P_{\delta_x} \text{ a.s.}$$

Proof. The inclusion

$$\{\mathcal{R}(Y^{(a)}) \cap A^c \neq \emptyset\} \supset \{\mathcal{R}(Y^{(a)}) \cap \bar{A}^c \neq \emptyset\}$$

is trivial. To show the reverse inclusion, we may argue as follows. By Theorem IV.9 in [**LG99**] (which holds under much less stringent assumptions on A), the event $\{\mathcal{R}(Y^{(a)}) \cap A^c \neq \emptyset\}$ holds if and only if the exit measure of the super-Brownian motion $Y^{(a)}$ from A is nonzero. Applying the special Markov property of superprocesses, we see that it is enough to prove that for super-Brownian motion starting from a nonzero initial measure supported on ∂A, the range immediately hits $(\bar{A})^c$. This is however easy under our regularity assumptions on A. We leave the details to the reader. ☐

From now on, we write $\{Y^{(a)}$ hits $F\}$ for the event $\{\mathcal{R}(Y^{(a)}) \cap F \neq \emptyset\}$.

3. Quenched convergence to super-Brownian motion

The main goal of this section is to prove Theorem 3.3. At the end of the section, we also establish Proposition 3.2, using certain arguments related to the proof of Theorem 3.3. To simplify notation, we set for every $\varpi \in \mathbf{\Omega}$ and $\varepsilon \in (0,1)$,

$$\Gamma^{\varepsilon}_{\varpi} = \sqrt{\varepsilon}\,\Gamma_{\varpi} = \bigcup_{i\in I} \sqrt{\varepsilon}(x_i + K_i).$$

The following lemma identifies a martingale problem solved by our branching Brownian motion $Z^{\varpi,\varepsilon}$. It can be proved by standard arguments: See e.g. Section 9.4 in [**EK86**].

LEMMA 3.9. *Let $\varpi \in \mathbf{\Omega}$ and $\varepsilon \geq 0$. Under each probability $\mathbb{P}_\mu$, $\mu \in \mathcal{M}_p(\mathbb{R}^d)$, the process $Z^{\varpi,\varepsilon}$ solves the following martingale problem: For every $f \in \bar{C}^2_+(\mathbb{R}^d)$ such that $0 < \inf f \leq f \leq 1$, the process*

$$e^{\langle Z^{\varpi,\varepsilon}_t, \log f\rangle} - e^{\langle Z^{\varpi,\varepsilon}_0, \log f\rangle} - \int_0^t \left\langle Z^{\varpi,\varepsilon}_s, \frac{\frac{1}{2}\Delta f + \varepsilon \mathbb{I}_{\Gamma_\varpi}(1-f) + \Upsilon(f) - f}{f}\right\rangle e^{\langle Z^{\varpi,\varepsilon}_s, \log f\rangle} ds$$

is a martingale.

We can derive from Lemma 3.9 (or from a direct argument) that for every $g \in \bar{C}^2(\mathbb{R}^d)$,

$$M_t(g) := \langle Z^{\varpi,\varepsilon}_t, g\rangle - \langle Z^{\varpi,\varepsilon}_0, g\rangle - \int_0^t \left\langle Z^{\varpi,\varepsilon}_s, \frac{1}{2}\Delta g - \varepsilon \mathbb{I}_{\Gamma_\varpi} g\right\rangle ds \tag{3.7}$$

is a martingale. An easy computation gives that the square bracket of this martingale is

$$[M(g), M(g)]_t = \int_0^t \langle Z^{\varpi,\varepsilon}_s, \nabla g.\nabla g\rangle\, ds + \sum_{0\leq s\leq t} \langle Z^{\varpi,\varepsilon}_s - Z^{\varpi,\varepsilon}_{s-}, g\rangle^2, \tag{3.8}$$

where the last sum in the right-hand side is an increasing process with compensator

$$\int_0^t \langle Z^{\varpi,\varepsilon}_s, (\sigma^2 + \varepsilon \mathbb{I}_{\Gamma_\varpi}) g^2\rangle\, ds. \tag{3.9}$$

The proof of Theorem 3.3 relies on the following two results, in which we use the notation $(\mathcal{Y}_t)_{t\geq 0}$ for the canonical process on $D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$.

LEMMA 3.10. *For every $\varpi \in \mathbf{\Omega}$ and $\varepsilon \in (0,1)$, let $\Pi^{\varpi,\varepsilon}$ be the law of the process $X^{\varpi,\varepsilon}$ under $\mathbb{P}_{[\varepsilon^{-1}]\delta_0}$. Then,*

(i) *For every $\delta > 0$ and $T > 0$, there is a compact subset $K_{\delta,T}$ of $\mathbb{R}^d$ such that, for every $\varpi \in \mathbf{\Omega}$,*

$$\sup_{\varepsilon\in(0,1)} \Pi^{\varpi,\varepsilon}\Big(\sup_{0\leq t\leq T} \mathcal{Y}_t(K^c_{\delta,T})\Big) < \delta$$

(ii) *For every* $g \in \bar{C}^2(\mathbb{R}^d)$ *and* $\varpi \in \mathbf{\Omega}$, *the collection of the laws of the process* $(\langle \mathcal{Y}_t, g\rangle)_{t\geq 0}$ *under* $\Pi^{\varpi,\varepsilon}$, $\varepsilon \in (0,1)$, *is relatively compact in the space of all probability measures on* $D_{\mathbb{R}}[0,\infty)$.

Consequently, for every $\varpi \in \mathbf{\Omega}$, *the collection* $(\Pi^{\varpi,\varepsilon})_{\varepsilon\in(0,1)}$ *is relatively compact in the space of all probability measures on* $D_{\mathcal{M}_f(\mathbb{R}^d)}[0,\infty)$.

The last assertion of the lemma is an immediate consequence of (i) and (ii) using Theorem II.4.1 in [**Per02**].

PROPOSITION 3.11. *Let* $g \in \bar{C}^2_+(\mathbb{R}^d)$. *There exists a measurable subset* $\mathbf{\Omega}_g$ *of* $\mathbf{\Omega}$ *such that* $\mathbf{P}(\mathbf{\Omega}_g) = 1$ *and the following holds for every* $\varpi \in \mathbf{\Omega}_g$. *For every* $s, t \geq 0$, *for every integer* $p \in \mathbb{N}$ *and every choice of* $t_1, \ldots, t_p \in [0,t]$ *and* $f_1, \ldots, f_p \in \bar{C}\big(\mathcal{M}_f(\mathbb{R}^d)\big)$, *we have*

$$\lim_{\varepsilon\to 0} \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\left[\left\{e^{-\langle X^{\varpi,\varepsilon}_{t+s},g\rangle} - e^{-\langle X^{\varpi,\varepsilon}_{t},g\rangle} - \int_t^{t+s}\left\langle X^{\varpi,\varepsilon}_u, -\frac{1}{2}\Delta g + \psi_{(\kappa)}(g)\right\rangle e^{-\langle X^{\varpi,\varepsilon}_{u},g\rangle}du\right\}\right.$$
$$\left.\times \prod_{i=1}^p f_i(X^{\varpi,\varepsilon}_{t_i})\right] = 0.$$

Theorem 3.3 is now a straightforward consequence of the preceding two statements. Indeed, choose a countable dense subset G of $\bar{C}^2_+(\mathbb{R}^d)$ and set

$$\mathbf{\Omega}' = \bigcap_{g\in G} \mathbf{\Omega}_g.$$

Fix $\varpi \in \mathbf{\Omega}'$. By Lemma 3.10, the collection $(\Pi^{\varpi,\varepsilon})_{\varepsilon\in(0,1)}$ is relatively compact. Let Π^* be a sequential limit of this collection as ε tends to 0. We deduce from Proposition 3.11 that, for every $g \in G$, for every $s, t \geq 0$ and every choice of $t_1, \ldots, t_p \in [0,t]$ and $f_1, \ldots, f_p \in \bar{C}\big(\mathcal{M}_f(\mathbb{R}^d)\big)$, we have

$$\Pi^*\left[\left\{e^{-\langle \mathcal{Y}_{t+s},g\rangle} - e^{-\langle \mathcal{Y}_{t},g\rangle} - \int_t^{t+s}\left\langle \mathcal{Y}_u, -\frac{1}{2}\Delta g + \psi_{(\kappa)}(g)\right\rangle e^{-\langle \mathcal{Y}_{u},g\rangle}du\right\}\prod_{i=1}^p f_i(\mathcal{Y}_{t_i})\right] = 0. \tag{3.10}$$

The required passage to the limit under the expectation sign is easily justified: Note that, by comparing with the case when there is no killing and using standard results of the theory of critical branching processes, we have for every $T > 0$

$$\sup_{\varepsilon\in(0,1)}\left(\mathbb{E}_{[\varepsilon^{-1}\delta_0]}\left[\sup_{0\leq r\leq T}\langle X^{\varpi,\varepsilon}_r, 1\rangle^2\right]\right) < \infty. \tag{3.11}$$

Since G is dense in $\bar{C}^2_+(\mathbb{R}^d)$, another easily justified passage to the limit shows that (3.10) holds for every $g \in \bar{C}^2_+(\mathbb{R}^d)$. Thus, Π^* satisfies the martingale problem for super-Brownian motion with branching mechanism $\psi_{(\kappa)}$ as stated in Section 2. Since it is also clear that $\Pi^*(Y_0 = \delta_0) = 1$, Π^* must be the law of super-Brownian motion with branching mechanism $\psi_{(\kappa)}$ started from δ_0. This completes the proof of Theorem 3.3, but we still need to establish Lemma 3.10 and Proposition 3.11. □

Proof of Lemma 3.10. The compact containment condition (i) in the lemma is immediately obtained by observing that $X^{\varpi,\varepsilon}$ is dominated by $X^0 = X^{\varpi,0}$ and by using the convergence of rescaled branching Brownian motions (without killing in the obstacles) towards super-Brownian motion. So we only need to verify (ii). In the remaining part of the proof, we fix $\varpi \in \mathbf{\Omega}$ and $g \in \bar{C}^2_+(\mathbb{R}^d)$. To simplify notation, we set

$$\mathcal{X}^\varepsilon_t = \langle X^{\varpi,\varepsilon}_t, g\rangle.$$

By the remarks following Lemma 3.9 and an elementary scaling transformation, we have

$$\mathcal{X}_t^\varepsilon = \mathcal{X}_0^\varepsilon + M_t^\varepsilon + V_t^\varepsilon,$$

where

$$V_t^\varepsilon = \int_0^t \Big\langle X_s^{\varpi,\varepsilon}, \frac{1}{2}\Delta g - \mathbb{I}_{\Gamma_\varpi^\varepsilon} g \Big\rangle \, ds$$

and M^ε is a martingale, whose square bracket is given by

$$[M^\varepsilon, M^\varepsilon]_t = \int_0^t \langle X_s^{\varpi,\varepsilon}, \varepsilon \nabla g . \nabla g\rangle \, ds + \sum_{0 \le s \le t} \langle X_s^{\varpi,\varepsilon} - X_{s-}^{\varpi,\varepsilon}, g\rangle^2.$$

Furthermore, the oblique bracket of M^ε is equal to

$$\langle M^\varepsilon, M^\varepsilon\rangle_t = \int_0^t \big\langle X_s^{\varpi,\varepsilon}, \sigma^2 g^2 + \varepsilon\big(\nabla g.\nabla g + \mathbb{I}_{\Gamma_\varpi^\varepsilon} g^2\big)\big\rangle \, ds.$$

By standard criteria (see in particular Theorem VI.4.13 in Jacod and Shiryaev [**JS87**]), the tightness of the laws of the processes $\mathcal{X}^\varepsilon$, $\varepsilon \in (0,1)$, will follow if we can verify that the laws of the processes V^ε and $\langle M^\varepsilon, M^\varepsilon\rangle$, for $\varepsilon \in (0,1)$, are C-tight. But this is immediate from the preceding explicit formulas and (3.11). □

Proof of Proposition 3.11. Let us fix $s, t \ge 0$, $t_1, \dots, t_p \in [0,t]$ and $f_1, \dots, f_p \in \bar{C}\big(\mathcal{M}_f(\mathbb{R}^d)\big)$. Let $\varpi \in \mathbf{\Omega}$ and $\varepsilon \in (0,1)$. Using the facts that $\Upsilon(1) = \Upsilon'(1) = 1$, $\Upsilon''(1) = \sigma^2$ and Lemma 3.9 applied to the function $f(x) = \exp\big\{-\varepsilon g(x\sqrt{\varepsilon})\big\}$, we can write

$$\begin{aligned}
0 = \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Bigg[\bigg\{ & e^{-\langle X_{t+s}^{\varpi,\varepsilon}, g\rangle} - e^{-\langle X_t^{\varpi,\varepsilon}, g\rangle} - \varepsilon^{-2}\int_t^{t+s} \Big\langle X_u^{\varpi,\varepsilon}, \frac{1}{2}\big(-\varepsilon^2\Delta g + \varepsilon^3 \nabla g.\nabla g\big) \\
& + \varepsilon \mathbb{I}_{\Gamma_\varpi^\varepsilon}\big(1 - e^{-\varepsilon g}\big)e^{\varepsilon g} + e^{\varepsilon g}\big(\Upsilon\big(e^{-\varepsilon g}\big) - e^{-\varepsilon g}\big)\Big\rangle \, e^{-\langle X_u^{\varpi,\varepsilon}, g\rangle} du\bigg\} \prod_{i=1}^p f_i\big(X_{t_i}^{\varpi,\varepsilon}\big)\Bigg] \\
= \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Bigg[\bigg\{ & e^{-\langle X_{t+s}^{\varpi,\varepsilon}, g\rangle} - e^{-\langle X_t^{\varpi,\varepsilon}, g\rangle} - \int_t^{t+s} \Big\langle X_u^{\varpi,\varepsilon}, -\frac{1}{2}\Delta g + \frac{\varepsilon}{2}\nabla g.\nabla g \\
& + \mathbb{I}_{\Gamma_\varpi^\varepsilon}(g + \eta_\varepsilon) + \frac{\sigma^2}{2}(g^2 + \tilde{\eta}_\varepsilon)\Big\rangle \, e^{-\langle X_u^{\varpi,\varepsilon}, g\rangle} du\bigg\} \prod_{i=1}^p f_i\big(X_{t_i}^{\varpi,\varepsilon}\big)\Bigg], \qquad (3.12)
\end{aligned}$$

where $\|\eta_\varepsilon\| \le C_1(g)\varepsilon$ for some constant $C_1(g)$ depending only on g, and $\|\tilde{\eta}_\varepsilon\| \to 0$ as $\varepsilon \to 0$. On the one hand, using Fubini's theorem, the fact that $(\langle X_r^{\varpi,\varepsilon}, 1\rangle)_{r \ge 0}$ is a supermartingale and the inequality $\mathbb{E}_{[\varepsilon^{-1}]\delta_0}[\langle X_0^{\varpi,\varepsilon}, 1\rangle] \le 1$, we have

$$\begin{aligned}
& \left| \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\left[\left\{ \int_t^{t+s} \Big\langle X_u^{\varpi,\varepsilon}, \frac{\varepsilon}{2}\nabla g.\nabla g + \mathbb{I}_{\Gamma_\varpi^\varepsilon}\eta_\varepsilon + \frac{\sigma^2}{2}\tilde{\eta}_\varepsilon \Big\rangle \, e^{-\langle X_u^{\varpi,\varepsilon}, g\rangle} du \right\} \prod_{i=1}^p f_i\big(X_{t_i}^{\varpi,\varepsilon}\big) \right] \right| \\
& \qquad \le \left(\frac{\varepsilon}{2}\big(\|\nabla g.\nabla g\| + 2C_1(g)\big) + \frac{\sigma^2}{2}\|\tilde{\eta}_\varepsilon\| \right) \left(\prod_{i=1}^p \|f_i\| \right) \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\left[\int_t^{t+s} \langle X_u^{\varpi,\varepsilon}, 1\rangle \, du \right] \\
& \qquad \le \big(C_2(g)\,\varepsilon + (\sigma^2/2)\|\tilde{\eta}_\varepsilon\|\big) \left(\prod_{i=1}^p \|f_i\| \right) s
\end{aligned}$$

with a constant $C_2(g)$ depending only on g. On the other hand,

$$\mathbb{E}_{[\varepsilon^{-1}]\delta_0}\left[\left\{ \int_t^{t+s} \langle X_u^{\varpi,\varepsilon}, \, \mathbb{I}_{\Gamma_\varpi^\varepsilon} g\rangle e^{-\langle X_u^{\varpi,\varepsilon}, g\rangle} du \right\} \prod_{i=1}^p f_i\big(X_{t_i}^{\varpi,\varepsilon}\big) \right]$$

$$
\begin{aligned}
&= \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\Big\{\int_t^{t+s} \langle X_u^{\varpi,\varepsilon}, \kappa g\rangle\, e^{-\langle X_u^{\varpi,\varepsilon}, g\rangle} du\Big\}\prod_{i=1}^p f_i\big(X_{t_i}^{\varpi,\varepsilon}\big)\Big]\\
&\quad + \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\Big\{\int_t^{t+s} \big\langle X_u^{\varpi,\varepsilon}, \big(\mathbb{I}_{\Gamma_\varpi^\varepsilon} - \kappa\big) g\big\rangle\, e^{-\langle X_u^{\varpi,\varepsilon}, g\rangle} du\Big\}\prod_{i=1}^p f_i\big(X_{t_i}^{\varpi,\varepsilon}\big)\Big]. \qquad (3.13)
\end{aligned}
$$

The absolute value of the second term in the right-hand side of (3.13) is bounded above by

$$
\Big(\prod_{i=1}^p \|f_i\|\Big)\int_t^{t+s} \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\big|\big\langle X_u^{\varpi,\varepsilon}, \big(\mathbb{I}_{\Gamma_\varpi^\varepsilon} - \kappa\big) g\big\rangle\big|\Big] du.
$$

Consequently, going back to (3.12) and using the preceding estimates, we obtain for every $\varpi \in \mathbf{\Omega}$ and $\varepsilon \in (0,1)$,

$$
\begin{aligned}
&\Bigg|\mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Bigg[\Big\{e^{-\langle X_{t+s}^{\varpi,\varepsilon}, g\rangle} - e^{-\langle X_t^{\varpi,\varepsilon}, g\rangle} - \int_t^{t+s}\Big\langle X_u^{\varpi,\varepsilon}, -\frac{1}{2}\Delta g + \kappa g + \frac{\sigma^2}{2} g^2\Big\rangle\, e^{-\langle X_u^{\varpi,\varepsilon}, g\rangle} du\Big\}\\
&\qquad\qquad \times \prod_{i=1}^p f_i\big(X_{t_i}^{\varpi,\varepsilon}\big)\Bigg]\Bigg|\\
&\qquad \le \Big(\prod_{i=1}^p \|f_i\|\Big)\Big(\big\{C_2(g)\,\varepsilon + (\sigma^2/2)\|\tilde{\eta}_\varepsilon\|\big\}s + r_\varepsilon(\varpi, g, t, t+s)\Big), \qquad (3.14)
\end{aligned}
$$

where $r_\varepsilon(\varpi, g, t_1, t_2)$ is defined for $0 \le t_1 \le t_2$ by

$$
r_\varepsilon(\varpi, g, t_1, t_2) = \int_{t_1}^{t_2} \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\big|\big\langle X_u^{\varpi,\varepsilon}, \big(\mathbb{I}_{\Gamma_\varpi^\varepsilon} - \kappa\big) g\big\rangle\big|\Big] du.
$$

LEMMA 3.12. *Let $u > 0$. Then, $\mathbf{P}$-a.s.*

$$
\lim_{\varepsilon\to\infty} \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\big|\big\langle X_u^{\varpi,\varepsilon},\ \big(\mathbb{I}_{\Gamma_\varpi^\varepsilon} - \kappa\big) g\big\rangle\big|\Big] = 0. \qquad (3.15)
$$

Assuming that the lemma is proved, we can readily complete the proof of Proposition 3.11. By Fubini's theorem, we have

$$
\begin{aligned}
\mathbf{P}\otimes\lambda_1&\Big[\Big\{(\varpi, u)\ :\ \limsup_{\varepsilon\to 0}\ \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\big|\big\langle X_u^{\varpi,\varepsilon},\ \big(\mathbb{I}_{\Gamma_\varpi^\varepsilon} - \kappa\big) g\big\rangle\big|\Big] > 0\Big\}\Big]\\
&= \int_0^\infty du\, \mathbf{P}\Big[\limsup_{\varepsilon\to 0}\ \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\big|\big\langle X_u^{\varpi,\varepsilon},\ \big(\mathbb{I}_{\Gamma_\varpi^\varepsilon} - \kappa\big) g\big\rangle\big|\Big] > 0\Big] = 0,
\end{aligned}
$$

by Lemma 3.12. Thus, there exists a set $\mathbf{\Omega}_g$ with $\mathbf{P}\big[\mathbf{\Omega}_g\big] = 1$ such that for every $\varpi \in \mathbf{\Omega}_g$, the convergence in (3.15) holds simultaneously for all $u > 0$, except possibly on a set of zero Lebesgue measure (depending on ϖ). Since for every $\varpi \in \mathbf{\Omega}$, $\varepsilon \in (0,1)$ and $u \ge 0$

$$
\mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\big|\big\langle X_u^{\varpi,\varepsilon},\ \big(\mathbb{I}_{\Gamma_\varpi^\varepsilon} - \kappa\big) g\big\rangle\big|\Big] \le \|g\|\,\mathbb{E}_{[\varepsilon^{-1}]\delta_0}[\langle X_u^{\varpi,\varepsilon}, 1\rangle] \le \|g\|,
$$

dominated convergence guarantees that for each $\varpi \in \mathbf{\Omega}_g$, $\lim_{\varepsilon\to 0} r_\varepsilon(\varpi, g, t, t+s) = 0$ for all $t, s \ge 0$. It follows that the right-hand side of (3.14) tends to 0 as $\varepsilon \to 0$ when $\varpi \in \mathbf{\Omega}_g$, which completes the proof of Proposition 3.11. □

Proof of Lemma 3.12. Recall that $\xi^{\varpi,\varepsilon}$ denotes standard d-dimensional Brownian motion killed at rate ε within Γ_ϖ. We also use the notation $\chi^{\varpi,\varepsilon}$ for Brownian motion killed at rate 1 within $\Gamma_\varpi^\varepsilon$. Both processes $\xi^{\varpi,\varepsilon}$ and $\chi^{\varpi,\varepsilon}$ start from x under the probability measure P_x.

We first recall classical moment formulas for branching Brownian motion. For every $x \in \mathbb{R}^d$, $k \in \mathbb{N}$, and every bounded measurable function h on $\mathbb{R}^d$, we have

$$
\mathbb{E}_{k\delta_x}[\langle Z_t^{\varpi,\varepsilon}, h\rangle] = k\,\mathrm{E}_x[h(\xi_t^{\varpi,\varepsilon})]
$$

and

$$\mathbb{E}_{k\delta_x}[\langle Z_t^{\varpi,\varepsilon}, h\rangle^2] = k\,\mathrm{E}_x[h(\xi_t^{\varpi,\varepsilon})^2] + k(k-1)\mathrm{E}_x[h(\xi_t^{\varpi,\varepsilon})]^2 + k\sigma^2\,\mathrm{E}_x\left[\int_0^t ds\,\mathrm{E}_{\xi_s^{\varpi,\varepsilon}}[h(\xi_{t-s}^{\varpi,\varepsilon})]^2\right].$$

These formulas are easily derived, first for $k = 1$, from the well-known formula for the Laplace functional of $Z_t^{\varpi,\varepsilon}$ under $\mathbb{P}_{\delta_x}$: See e.g. Proposition II.3 in [**LG99**].

Recalling the definition of $X^{\varpi,\varepsilon}$ in terms of $Z^{\varpi,\varepsilon}$, we get similar formulas for the first and second moment of $\langle X_s^{\varpi,\varepsilon}, h\rangle$. In particular, for $s \geq 0$ and for every $\varpi \in \mathbf{\Omega}$ and $\varepsilon \in (0,1)$, we have

$$\text{(3.16)}\quad \mathbb{E}_{[\varepsilon^{-1}]\delta_0}\Big[\big\langle X_s^{\varpi,\varepsilon}, \big(\mathbb{I}_{\Gamma_\varpi^\varepsilon} - \kappa\big)g\big\rangle^2\Big] = \varepsilon^2\big([\varepsilon^{-1}]^2 - [\varepsilon^{-1}]\big)A_1^{\varpi,\varepsilon}(0,s,g)^2 + \varepsilon[\varepsilon^{-1}]A_2^{\varpi,\varepsilon}(0,s,g),$$

where, for every $x \in \mathbb{R}^d$,

$$A_1^{\varpi,\varepsilon}(x,s,g) := \mathrm{E}_x\Big[\big(\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\chi_s^{\varpi,\varepsilon}) - \kappa\big)g(\chi_s^{\varpi,\varepsilon})\Big],$$

and

$$A_2^{\varpi,\varepsilon}(x,s,g) := \sigma^2\,\mathrm{E}_x\left[\int_0^s A_1^{\varpi,\varepsilon}(\chi_v^{\varpi,\varepsilon}, s-v, g)^2 dv\right] + \varepsilon\,\mathrm{E}_x\Big[\big(\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\chi_s^{\varpi,\varepsilon}) - \kappa\big)^2 g(\chi_s^{\varpi,\varepsilon})^2\Big].$$

We claim that, for every $x \in \mathbb{R}^d$ and $s > 0$,

$$\text{(3.17)}\qquad \mathbf{P}\text{–a.s.},\qquad \lim_{\varepsilon\to 0}\mathrm{E}_x\Big[\big(\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\chi_s^{\varpi,\varepsilon}) - \kappa\big)g(\chi_s^{\varpi,\varepsilon})\Big] = 0.$$

Assume for the moment that the claim holds. By taking $x = 0$ in (3.17), we obtain that $A_1^{\varpi,\varepsilon}(0,s,g)$ tends to 0 as $\varepsilon \to 0$, $\mathbf{P}$-a.s. Consider next $A_2^{\varpi,\varepsilon}(0,s,g)$. The second term in the formula for $A_2^{\varpi,\varepsilon}(x,s,g)$ obviously tends to 0 uniformly in x, s and independently of the obstacles, as $\varepsilon \to 0$. To handle the first term, use Fubini's theorem to obtain that, for every ϖ belonging to a set $\widetilde{\mathbf{\Omega}}$ of full probability, there exists a set $\mathcal{N}_\varpi \subset \mathbb{R}^d \times \mathbb{R}_+$ of zero Lebesgue measure such that the convergence in (3.17) holds simultaneously for all $(x,s) \notin \mathcal{N}_\varpi$. We can then write the first term in the formula for $A_2^{\varpi,\varepsilon}(0,s,g)$ as follows:

$$\text{(3.18)}\qquad \sigma^2\int_0^s dv\int_{\mathbb{R}^d} dy\,p_v^{\varpi,\varepsilon}(0,y)\,\mathrm{E}_y\Big[\big(\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\chi_{s-v}^{\varpi,\varepsilon}) - \kappa\big)g(\chi_{s-v}^{\varpi,\varepsilon})\Big]^2,$$

where $p_v^{\varpi,\varepsilon}(\cdot,\cdot)$ is the transition kernel of $\chi^{\varpi,\varepsilon}$ at time v. Plainly, we have $p_v^{\varpi,\varepsilon}(x,y) \leq p_v(x,y)$ for each $v \geq 0$ and $\varepsilon \in (0,1)$, where $p_v(\cdot,\cdot)$ is the transition kernel of standard Brownian motion at time v. Using this bound and dominated convergence gives us that the quantity in (3.18) tends to 0 as $\varepsilon \to 0$, for every $\varpi \in \widetilde{\mathbf{\Omega}}$. Hence $A_2^{\varpi,\varepsilon}(0,s,g)$ tends to 0 as $\varepsilon \to 0$, $\mathbf{P}$-a.s. The result of Lemma 3.12 now follows from (3.16) and the Cauchy-Schwarz inequality.

We still have to prove our claim (3.17). We fix $x \in \mathbb{R}^d$ and $s > 0$. Let $\varpi \in \mathbf{\Omega}$ and $\varepsilon \in (0,1)$. From the definition of Brownian motion killed at rate 1 in $\Gamma_\varpi^\varepsilon$, we have

$$\text{(3.19)}\quad \mathrm{E}_x\Big[\big(\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\chi_s^{\varpi,\varepsilon}) - \kappa\big)g(\chi_s^{\varpi,\varepsilon})\Big] = \mathrm{E}_x\Big[\big(\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\xi_s) - \kappa\big)g(\xi_s)\exp\Big\{-\int_0^s \mathbb{I}_{\Gamma_\varpi^\varepsilon}(\xi_u)du\Big\}\Big].$$

Let $\eta > 0$ and choose $\theta \in (0,s)$ such that $e^{-\theta} \geq 1 - \eta$. By the Markov property applied at time $s - \theta$, the quantities in (3.19) are equal to

$$\begin{aligned}
&\mathrm{E}_x\left[\exp\Big\{-\int_0^{s-\theta}\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\xi_u)du\Big\}\,\mathrm{E}_{\xi_{s-\theta}}\Big[\big(\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\xi_\theta) - \kappa\big)g(\xi_\theta)\exp\Big\{-\int_0^\theta \mathbb{I}_{\Gamma_\varpi^\varepsilon}(\xi_v)dv\Big\}\Big]\right] \\
&\quad = \mathrm{E}_x\Big[\exp\Big\{-\int_0^{s-\theta}\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\xi_u)du\Big\}\,\mathrm{E}_{\xi_{s-\theta}}\Big[\big(\mathbb{I}_{\Gamma_\varpi^\varepsilon}(\xi_\theta) - \kappa\big)g(\xi_\theta)\Big]\Big]
\end{aligned}$$

$$+\mathrm{E}_x\Big[\exp\Big\{-\int_0^{s-\theta}\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_u)du\Big\}\,\mathrm{E}_{\xi_{s-\theta}}\Big[\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_\theta)-\kappa\big)g(\xi_\theta)$$
$$\times\Big(\exp\Big\{-\int_0^{\theta}\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_v)dv\Big\}-1\Big)\Big]\Big]. \tag{3.20}$$

The condition $1-e^{-\theta}\le\eta$ entails that the absolute value of the second term in the right-hand side of (3.20) is bounded above by $\|g\|\eta$. Suppose we know that for every $y\in\mathbb{R}^d$ we have

$$\mathbf{P}\text{−a.s.,}\qquad \lim_{\varepsilon\to0}\mathrm{E}_y\Big[\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_\theta)-\kappa\big)g(\xi_\theta)\Big]=0. \tag{3.21}$$

Then, using the fact that the law of $\xi_{s-\theta}$ is absolutely continuous with respect to Lebesgue measure (and Fubini's theorem to see that the convergence in (3.21) holds simultaneously for almost all $y\in\mathbb{R}^d$, $\mathbf{P}$-a.s.), we conclude that the first term in the right-hand side of (3.20) converges to 0 with $\mathbf{P}$-probability 1 as $\varepsilon\to0$. Consequently, we have $\mathbf{P}$-a.s.

$$\limsup_{\varepsilon\to0}\ \mathrm{E}_x\Big[\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\chi^{\varpi,\varepsilon}_s)-\kappa\big)g(\chi^{\varpi,\varepsilon}_s)\Big]\le\|g\|\eta.$$

Since η was arbitrary, our claim (3.17) follows.

It thus remains to prove that (3.21) holds. We fix $y\in\mathbb{R}^d$ and $\theta>0$. In the following, ξ' stands for another Brownian motion independent of ξ, which also starts from y under the probability measure P_y. For each $\varepsilon\in(0,1)$, we have

$$\begin{aligned}\mathbf{E}\Big[\mathrm{E}_y\big[\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_\theta)-\kappa\big)g(\xi_\theta)\big]^2\Big]&=\mathbf{E}\Big[\mathrm{E}_y\big[\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_\theta)-\kappa\big)\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi'_\theta)-\kappa\big)g(\xi_\theta)g(\xi'_\theta)\big]\Big]\\&=\mathrm{E}_y\Big[g(\xi_\theta)g(\xi'_\theta)\Big\{\mathbf{P}\big[\xi_\theta\in\Gamma^\varepsilon_\varpi;\ \xi'_\theta\in\Gamma^\varepsilon_\varpi\big]-\kappa^2\Big\}\Big],\end{aligned}\tag{3.22}$$

where the last line uses Fubini's theorem and the definition of κ. Recall that the measure Θ is supported on compact sets which are contained in the fixed ball $\overline{B}(0,r_0)$. If $|x-x'|>2r_0$, the sets $\{(z,K):x\in z+K\text{ and }K\subset\overline{B}(0,r_0)\}$ and $\{(z,K):x'\in z+K\text{ and }K\subset\overline{B}(0,r_0)\}$ are disjoint, and so the events $\{x\in\Gamma_\varpi\}$ and $\{x'\in\Gamma_\varpi\}$ are independent. Recalling that $\Gamma^\varepsilon_\varpi=\sqrt{\varepsilon}\,\Gamma_\varpi$, we see that if $|x-x'|>2r_0\sqrt{\varepsilon}$ we have $\mathbf{P}[x\in\Gamma^\varepsilon_\varpi;\ x'\in\Gamma^\varepsilon_\varpi]=\mathbf{P}[x\in\Gamma^\varepsilon_\varpi]\mathbf{P}[x'\in\Gamma^\varepsilon_\varpi]=\kappa^2$, which enables us to write

$$\mathbf{P}\big[\xi_\theta\in\Gamma^\varepsilon_\varpi;\ \xi'_\theta\in\Gamma^\varepsilon_\varpi\big]=\kappa^2\,\mathbb{I}_{\{|\xi_\theta-\xi'_\theta|>2r_0\sqrt{\varepsilon}\}}+\mathbf{P}\big[\xi_\theta\in\Gamma^\varepsilon_\varpi;\ \xi'_\theta\in\Gamma^\varepsilon_\varpi\big]\,\mathbb{I}_{\{|\xi_\theta-\xi'_\theta|\le2r_0\sqrt{\varepsilon}\}}.$$

Going back to (3.22), we obtain

$$\begin{aligned}\mathbf{E}\Big[\mathrm{E}_y\big[\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_\theta)-\kappa\big)g(\xi_\theta)\big]^2\Big]&=\mathrm{E}_y\Big[g(\xi_\theta)g(\xi'_\theta)\mathbb{I}_{\{|\xi_\theta-\xi'_\theta|\le2r_0\sqrt{\varepsilon}\}}\Big\{\mathbf{P}\big[\xi_\theta\in\Gamma^\varepsilon_\varpi;\ \xi'_\theta\in\Gamma^\varepsilon_\varpi\big]-\kappa^2\Big\}\Big]\\&\le\|g\|^2\mathrm{P}_y\big[|\xi_\theta-\xi'_\theta|\le2r_0\sqrt{\varepsilon}\big]\\&=\|g\|^2\mathrm{P}_0\big[|\xi_{2\theta}|\le2r_0\sqrt{\varepsilon}\big]\\&\le C\frac{\|g\|^2}{\theta^{d/2}}\,\varepsilon^{d/2},\end{aligned}\tag{3.23}$$

where the constant $C>0$ depends only on r_0. Let $\eta>0$. By the Markov inequality, we can write

$$\mathbf{P}\Big[\big|\mathrm{E}_y\big[\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_\theta)-\kappa\big)g(\xi_\theta)\big]\big|>\eta\Big]\le\frac{1}{\eta^2}\,\mathbf{E}\Big[\mathrm{E}_y\big[\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_\theta)-\kappa\big)g(\xi_\theta)\big]^2\Big]\le C\frac{\|g\|^2}{\eta^2\theta^{d/2}}\,\varepsilon^{d/2}.$$

Applying the last bound with $\varepsilon=\varepsilon_k=k^{-3}$ for every $k\in\mathbb{N}$, yields a convergent series, and so by the Borel-Cantelli lemma we have $\mathbf{P}$-a.s.,

$$\limsup_{k\to\infty}\ \Big|\mathrm{E}_y\big[\big(\mathbb{I}_{\Gamma^{\varepsilon_k}_\varpi}(\xi_\theta)-\kappa\big)g(\xi_\theta)\big]\Big|\le\eta.$$

Since this is satisfied for every choice of $\eta > 0$, the convergence (3.21) holds along the sequence $(\varepsilon_k)_{k\geq 1}$.

To complete the proof, we set for every $\varpi \in \mathbf{\Omega}$ and $\varepsilon \in (0,1)$

$$U_{\varpi,\varepsilon} = \mathrm{E}_y\big[\mathbb{I}_{\Gamma^\varepsilon_\varpi}(\xi_\theta)\, g(\xi_\theta)\big]$$

(recall that y and θ are fixed). We shall prove the following result: For every $\varpi \in \mathbf{\Omega}$,

$$\lim_{k\to\infty} \sup_{(k+1)^{-3}\leq\varepsilon\leq k^{-3}} \big|U_{\varpi,\varepsilon} - U_{\varpi,\varepsilon_k}\big| = 0. \tag{3.24}$$

Combining (3.24) with the fact that the convergence (3.21) is true along the sequence $(\varepsilon_k)_{k\geq 1}$ will then lead to the desired conclusion. In order to prove (3.24), we first note that for every $\varepsilon \in (0,1)$,

$$U_{\varpi,\varepsilon} = \frac{1}{(2\pi\theta)^{d/2}} \int_{\mathbb{R}^d} \mathbb{I}_{\Gamma^\varepsilon_\varpi}(x)\, g(x) e^{-\frac{|x-y|^2}{2\theta^2}}\, dx = \int_{\mathbb{R}^d} \mathbb{I}_{\Gamma^\varepsilon_\varpi}(x)\, h(x) dx,$$

with a function $h \in \bar{C}^2_+(\mathbb{R}^d)$ which depends on y and θ, but not on ϖ. Furthermore, for any fixed $\eta > 0$ we can find a large closed ball B centered at the origin and such that $\int_{B^c} h(x)dx < \eta$. Hence, if we set

$$U'_{\varpi,\varepsilon} = \int_B h(x)\, \mathbb{I}_{\Gamma^\varepsilon_\varpi}(x) dx,$$

we have $|U_{\varpi,\varepsilon} - U'_{\varpi,\varepsilon}| \leq \eta$ for every $\varepsilon \in (0,1)$. Thanks to this remark, it is enough to prove that (3.24) holds when $U_{\varpi,\varepsilon}$ and U_{ϖ,ε_k} are replaced by $U'_{\varpi,\varepsilon}$ and $U'_{\varpi,\varepsilon_k}$ respectively.

Let $k \in \mathbb{N}$ and $\varepsilon \in [(k+1)^{-3}, k^{-3})$. We have

$$\begin{aligned} U'_{\varpi,\varepsilon} &= \varepsilon^{d/2} \int_{\varepsilon^{-1/2}B} h\big(x\sqrt{\varepsilon}\big)\mathbb{I}_{\Gamma_\varpi}(x) dx \\ &= \varepsilon^{d/2} \int_{k^{3/2}B} h\big(x\sqrt{\varepsilon}\big)\mathbb{I}_{\Gamma_\varpi}(x) dx + \varepsilon^{d/2} \int_{(\varepsilon^{-1/2}B)\setminus(k^{3/2}B)} h\big(x\sqrt{\varepsilon}\big)\mathbb{I}_{\Gamma_\varpi}(x) dx. \end{aligned} \tag{3.25}$$

Now, the first term in the right-hand side of (3.25) is equal to

$$\begin{aligned} &\big(\varepsilon k^3\big)^{d/2} \frac{1}{k^{3d/2}} \int_{k^{3/2}B} h\Big(\frac{x}{k^{3/2}}\Big)\mathbb{I}_{\Gamma_\varpi}(x) dx + \varepsilon^{d/2} \int_{k^{3/2}B} \Big[h\Big(x\sqrt{\varepsilon}\Big) - h\Big(\frac{x}{k^{3/2}}\Big)\Big]\mathbb{I}_{\Gamma_\varpi}(x) dx \\ &= \big(\varepsilon k^3\big)^{d/2} U'_{\varpi,\varepsilon_k} + \iota(\varpi,\varepsilon,k), \end{aligned}$$

where

$$\iota(\varpi,\varepsilon,k) = (\varepsilon k^3)^{d/2} \int_B \Big[h((k^{3/2}\sqrt{\varepsilon})x) - h(x)\Big]\, \mathbb{I}_{\Gamma_\varpi}(x) dx.$$

Note that $0 \leq 1 - k^{3/2}\sqrt{\varepsilon} \leq \frac{C}{k}$ with a constant C independent of ε and k, from which it easily follows that

$$\sup_{(k+1)^{-3}\leq\varepsilon\leq k^{-3}} |\iota(\varpi,\varepsilon,k)| \leq \frac{C'}{k}\, \|\nabla h\|,$$

with a constant C' depending only on B.

Similarly, the second term in the right-hand side of (3.25) is bounded above by

$$\varepsilon^{d/2}\, \|h\| \big[\varepsilon^{-d/2} - k^{3d/2}\big] \lambda_d(B) \leq \frac{C''\|h\|}{k}.$$

Finally, from (3.25) and the preceding estimates, we have

$$\big|U'_{\varpi,\varepsilon} - U'_{\varpi,\varepsilon_k}\big| \leq (1 - (\varepsilon k^3)^{d/2})\, U'_{\varpi,\varepsilon_k} + \frac{C'\|\nabla h\|}{k} + \frac{C''\|h\|}{k},$$

and the convergence (3.24) follows. This completes the proof. $\square$

Proof of Proposition 3.2. Clearly we get a lower bound for the quantity $\mathbb{P}_{\delta_0}(Z^{\varpi,\varepsilon} \text{ hits } B(0,R)^c)$ if we replace $Z^{\varpi,\varepsilon}$ by branching Brownian motion killed homogeneously over $\mathbb{R}^d$ at rate ε. Part (i) of the proposition thus follows from the lower bound in Lemma 3.5.

Let us turn to the proof of part (ii). We have the following inequality: For every $\varpi \in \mathbf{\Omega}$ and $\varepsilon \in (0,1)$,

$$\mathbb{P}_{\delta_0}\big[Z^{\varpi,\varepsilon} \text{ hits } B(0,R)^c\big] \leq \mathrm{P}_0\big[\xi^{\varpi,\varepsilon} \text{ hits } B(0,R)^c\big]. \tag{3.26}$$

Indeed, using the formalism of Subsection 2.3, the criticality of the offspring distribution can be used to check that the right-hand side of (3.26) is just the expected value of the number of those historical paths ω^v that first exit $B(0,R)$ during the interval $[\alpha_v, \beta_v \wedge \zeta_v^{\varpi,\varepsilon})$. Alternatively, it is easy to derive an integral equation similar to (3.6) for the function $x \to \mathbb{P}_{\delta_x}\big[Z^{\varpi,\varepsilon} \text{ hits } B(0,R)^c\big]$, and the bound (3.26) then trivially follows from this integral equation.

Then, let us bound $\mathbf{P} \otimes \mathrm{P}_0\big[\xi^{\varpi,\varepsilon} \text{ hits } B(0,R)^c\big]$. We write $\xi^{\varpi,\varepsilon,i}$ for the i-th coordinate of $\xi^{\varpi,\varepsilon}$, for every $i = 1, \ldots, d$. First, observe that

$$\mathbf{P} \otimes \mathrm{P}_0\big[\xi^{\varpi,\varepsilon} \text{ hits } B(0,R)^c\big] \leq \sum_{i=1}^{d} \mathbf{P} \otimes \mathrm{P}_0\big[\xi^{\varpi,\varepsilon,i} \text{ hits } (-R/\sqrt{d}, R/\sqrt{d})^c\big]. \tag{3.27}$$

Clearly, we can restrict our attention to the first term in the sum. Define $N_{\varepsilon,R} := [R\sqrt{\varepsilon}/\sqrt{d}]$ and for every $i \in \mathbb{Z}$, set $T_i^\varepsilon = \inf\{t \geq 0 : \xi_t^1 = i\varepsilon^{-1/2}\}$, where ξ^1 stands for the first coordinate of ξ. In this notation, we can write

$$\begin{aligned}
&\mathbf{P} \otimes \mathrm{P}_0\big[\xi^{\varpi,\varepsilon,1} \text{ hits } (-R/\sqrt{d}, R/\sqrt{d})^c\big] \\
&\qquad \leq \mathbf{P} \otimes \mathrm{P}_0\Big[\xi^{\varpi,\varepsilon,1} \text{ hits } \big(-N_{\varepsilon,R}\varepsilon^{-1/2}, N_{\varepsilon,R}\varepsilon^{-1/2}\big)^c\Big] \\
&\leq \mathbf{E} \otimes \mathrm{E}_0\bigg[\exp\bigg\{-\varepsilon \int_0^{T^\varepsilon_{N_{\varepsilon,R}}} \mathbb{I}_{\Gamma_\varpi}(\xi_s)ds\bigg\}\bigg] + \mathbf{E} \otimes \mathrm{E}_0\bigg[\exp\bigg\{-\varepsilon \int_0^{T^\varepsilon_{-N_{\varepsilon,R}}} \mathbb{I}_{\Gamma_\varpi}(\xi_s)ds\bigg\}\bigg].
\end{aligned} \tag{3.28}$$

Consider the first term in the right-hand side of (3.28). Set for every $i \in \mathbb{N}$ and $\varepsilon > 0$,

$$I_i^\varepsilon := \big\{j \in I : \ x_j \in ((i-1)\varepsilon^{-1/2}, i\varepsilon^{-1/2}) \times \mathbb{R}^{d-1}\big\},$$
$$\Gamma_\varpi(i,\varepsilon) := \bigcup_{j \in I_i^\varepsilon} (x_j + K_j) \qquad \text{and} \qquad \Gamma_\varpi^\varepsilon(i) := \sqrt{\varepsilon}\,\Gamma_\varpi(i,\varepsilon).$$

Note that the random sets $\Gamma_\varpi(i,\varepsilon)$, $i \in \mathbb{N}$, are independent under $\mathbf{P}$, by properties of Poisson measures. We have then

$$\begin{aligned}
\mathbf{E} \otimes \mathrm{E}_0\bigg[\exp\bigg\{-\varepsilon \int_0^{T^\varepsilon_{N_{\varepsilon,R}}} \mathbb{I}_{\Gamma_\varpi}(\xi_s)ds\bigg\}\bigg] &\leq \mathbf{E} \otimes \mathrm{E}_0\bigg[\exp\bigg\{-\varepsilon \sum_{i=1}^{N_{\varepsilon,R}} \int_{T^\varepsilon_{i-1}}^{T_i^\varepsilon} \mathbb{I}_{\Gamma_\varpi(i,\varepsilon)}(\xi_s)ds\bigg\}\bigg] \\
&= \mathbf{E} \otimes \mathrm{E}_0\bigg[\exp\bigg\{-\varepsilon \int_0^{T_1^\varepsilon} \mathbb{I}_{\Gamma_\varpi(1,\varepsilon)}(\xi_s)ds\bigg\}\bigg]^{N_{\varepsilon,R}},
\end{aligned} \tag{3.29}$$

where the equality comes from an application of the strong Markov property of ξ, together with the independence of the random sets $\Gamma_\varpi(i,\varepsilon)$ and the fact that the distribution of each of these random sets is invariant under translations by elements of $\{0\} \times \mathbb{R}^{d-1}$. By scaling, if T_1 denotes the entrance time of ξ into $[1,\infty) \times \mathbb{R}^{d-1}$, we can write

$$\alpha_\varepsilon := \mathbf{E} \otimes \mathrm{E}_0\bigg[\exp\bigg\{-\varepsilon \int_0^{T_1^\varepsilon} \mathbb{I}_{\Gamma_\varpi(1,\varepsilon)}(\xi_s)ds\bigg\}\bigg] = \mathbf{E} \otimes \mathrm{E}_0\bigg[\exp\bigg\{-\int_0^{T_1} \mathbb{I}_{\Gamma^\varepsilon_\varpi(1)}(\xi_s)ds\bigg\}\bigg].$$

We then observe that

$$\int_0^{T_1} \mathbb{I}_{\Gamma^\varepsilon_\varpi(1)}(\xi_s)\mathbb{I}_{(0,1)}(\xi^1_s)\, ds \underset{\varepsilon\to 0}{\overset{(P)}{\longrightarrow}} \kappa \int_0^{T_1} \mathbb{I}_{(0,1)}(\xi^1_s) ds \tag{3.30}$$

where the notation $\overset{(P)}{\longrightarrow}$ refers to convergence in probability under $\mathbf{P}\otimes \mathrm{P}_0$. To see this, we use arguments similar to the proof of Lemma 3.12. Notice that $\mathbf{P}[y\in\Gamma^\varepsilon_\varpi(1)]\leq\kappa$ if $y\in(0,1)\times\mathbb{R}^{d-1}$, with equality if $y\in(r_0\sqrt{\varepsilon},1-r_0\sqrt{\varepsilon})\times\mathbb{R}^{d-1}$. Using this remark, and the same argument as in the derivation of (3.23), we can write for every fixed $u>0$,

$$\begin{aligned}
&\mathbf{E}\otimes\mathrm{E}_0\bigg[\bigg\{\int_0^{T_1\wedge u}\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi(1)}(\xi_s)-\kappa\big)\mathbb{I}_{(0,1)}(\xi^1_s)ds\bigg\}^2\bigg]\\
&=\mathbf{E}\otimes\mathrm{E}_0\bigg[\int_0^{T_1\wedge u}\int_0^{T_1\wedge u}\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi(1)}(\xi_s)-\kappa\big)\big(\mathbb{I}_{\Gamma^\varepsilon_\varpi(1)}(\xi_t)-\kappa\big)\,\mathbb{I}_{(0,1)}(\xi^1_s)\,\mathbb{I}_{(0,1)}(\xi^1_t)\,dsdt\bigg]\\
&=\mathrm{E}_0\bigg[\int_0^{T_1\wedge u}\int_0^{T_1\wedge u}\big\{\mathbf{P}\big[\xi_s,\xi_t\in\Gamma^\varepsilon_\varpi(1)\big]-\kappa^2\big\}\,\mathbb{I}_{(0,1)}(\xi^1_s)\,\mathbb{I}_{(0,1)}(\xi^1_t)\,dsdt\bigg]+\mathcal{O}(\varepsilon^{1/2})\\
&=\mathrm{E}_0\bigg[\int_0^{T_1\wedge u}\int_0^{T_1\wedge u}\big\{\mathbf{P}\big[\xi_s,\xi_t\in\Gamma^\varepsilon_\varpi(1)\big]-\kappa^2\big\}\,\mathbb{I}_{\{|\xi_s-\xi_t|\leq 2r_0\sqrt{\varepsilon}\}}\,\mathbb{I}_{(0,1)}(\xi^1_s)\,\mathbb{I}_{(0,1)}(\xi^1_t)\,dsdt\bigg]\\
&\qquad\qquad+\mathcal{O}(\varepsilon^{1/2})\\
&\leq\mathrm{E}_0\bigg[\int_0^u\int_0^u\mathbb{I}_{\{|\xi_s-\xi_t|\leq 2r_0\sqrt{\varepsilon}\}}\,dsdt\bigg]+\mathcal{O}(\varepsilon^{1/2})
\end{aligned}$$

where the error term $\mathcal{O}(\varepsilon^{1/2})$ corresponds to the contribution of times s,t such that ξ^1_s or ξ^1_t belongs to the set $(0,r_0\sqrt{\varepsilon}]\cup[1-r_0\sqrt{\varepsilon},1)$. The preceding quantity tends to 0 as $\varepsilon\to 0$, which yields the convergence (3.30). Since the limiting variable in (3.30) is (strictly) positive a.s., we can find $\varepsilon_1\in(0,1)$ and $c_1<1$ such that $\alpha_\varepsilon\leq c_1$ for every $\varepsilon\in(0,\varepsilon_1)$. Using (3.27), (3.28) and (3.29), we arrive at

$$\mathbf{P}\otimes\mathrm{P}_0\big[\xi^{\varpi,\varepsilon}\text{ hits }B(0,R)^c\big]\leq 2d\,c_1^{N_{\varepsilon,R}},$$

for $\varepsilon\in(0,\varepsilon_1)$. This completes the proof of the upper bound of Proposition 3.2. □

4. Proof of the main result

In this section, we prove Theorem 3.1. We fix the environment ϖ such that the weak convergence of Theorem 3.3 holds, and derive the convergence in Theorem 3.1 for this fixed value of the environment. For the sake of simplicity, we shall omit ϖ in the notation and write Z^ε instead of $Z^{\varpi,\varepsilon}$, and X^ε instead of $X^{\varpi,\varepsilon}$.

We shall verify that for any increasing sequence $(R_n)_{n\geq 1}$ of positive reals converging to $+\infty$ and any sequence $(\varepsilon_n)_{n\geq 1}$ of nonnegative reals such that $\varepsilon_n R_n^2\to a\in[0,\infty]$, we have

$$\lim_{n\to\infty} R_n^2\,\mathbb{P}_{\delta_0}(Z^{\varepsilon_n}\text{ hits }R_nA^c)=u_{(\kappa a)}(0), \tag{3.31}$$

where $u_{(\infty)}(0)=0$ by convention.

The statement of Theorem 3.1 follows from this convergence. Indeed, if the conclusion of the theorem fails, then we can find a sequence $R_n\uparrow\infty$ and a sequence (ε_n) of nonnegative reals such that, for every $n\geq 1$,

$$|R_n^2\,\mathbb{P}_{\delta_0}(Z^{\varepsilon_n}\text{ hits }R_nA^c)-u_{(\kappa\varepsilon_nR_n^2)}(0)|\geq\delta$$

for some constant $\delta>0$. By extracting a subsequence, we may assume that $\varepsilon_nR_n^2\to a\in[0,\infty]$ and thus obtain a contradiction with (3.31) since we know from Lemma 3.7 that the mapping $b\to u_{(b)}(0)$ is continuous on $[0,\infty]$.

In proving (3.31), we may assume that $\varepsilon_n \to 0$ as $n \to \infty$. Indeed, suppose that (3.31) holds in this particular case and let (ε'_n) be a sequence that does not converge to 0. If the sequence $\varepsilon'_n R_n^2$ converges then necessarily its limit is $+\infty$, and we can find another sequence ε''_n such that $0 \le \varepsilon''_n \le \varepsilon'_n$, $\varepsilon''_n \to 0$ and $\varepsilon''_n R_n^2 \to \infty$. So, if we know that (3.31) holds in the case when the sequence (ε_n) tends to 0, we obtain

$$\lim_{n\to\infty} R_n^2 \, \mathbb{P}_{\delta_0}(Z^{\varepsilon''_n} \text{ hits } R_n A^c) = 0.$$

However, from the inequality $\varepsilon''_n \le \varepsilon'_n$ and a coupling argument (obvious if one uses the construction described in Subsection 2.3), we have

$$R_n^2 \, \mathbb{P}_{\delta_0}(Z^{\varepsilon'_n} \text{ hits } R_n A^c) \le R_n^2 \, \mathbb{P}_{\delta_0}(Z^{\varepsilon''_n} \text{ hits } R_n A^c)$$

giving the desired result for the sequence (ε'_n).

A similar comparison argument shows that it is enough to prove (3.31) in the case when $a < \infty$. Otherwise, it suffices to replace ε_n by $\varepsilon_n \wedge bR_n^{-2}$ and let $b \to \infty$, using the fact that $u_{(b)}(0) \to 0$ as $b \to \infty$.

Let us now proceed to the proof of (3.31). We fix the sequences $R_n \uparrow \infty$ and $\varepsilon_n \to 0$ such that $\varepsilon_n R_n^2 \to a \in [0,\infty)$. We first assume that $a > 0$. The case $a = 0$ will be discussed at the end of the section.

Let B be a closed subset of $\mathbb{R}^d$. For every $\varepsilon > 0$, we have by the definition of X^ε

$$\begin{aligned}
\mathbb{P}_{[\varepsilon^{-1}]\delta_0}(X^\varepsilon \text{ hits } B) &= \mathbb{P}_{[\varepsilon^{-1}]\delta_0}(\exists t \ge 0 : X_t^\varepsilon(B) > 0) \\
&= \mathbb{P}_{[\varepsilon^{-1}]\delta_0}\Big(\exists t \ge 0 : \int Z^\varepsilon_{\varepsilon^{-1}t}(dx)\, \mathbb{I}_B(x\sqrt{\varepsilon}) > 0\Big) \\
&= \mathbb{P}_{[\varepsilon^{-1}]\delta_0}(Z^\varepsilon \text{ hits } \varepsilon^{-1/2}B) \\
&= 1 - \mathbb{P}_{\delta_0}(Z^\varepsilon \text{ does not hit } \varepsilon^{-1/2}B)^{[\varepsilon^{-1}]},
\end{aligned}$$

since (for a fixed environment) the law of Z^ε under $\mathbb{P}_{[\varepsilon^{-1}]\delta_0}$ is obtained by adding $[\varepsilon^{-1}]$ independent copies of Z^ε under $\mathbb{P}_{\delta_0}$. Applying the preceding identity with $\varepsilon = \varepsilon_n$ and $B = b_n A^c$, where $b_n = \varepsilon_n^{1/2} R_n$, gives us that

$$1 - \mathbb{P}_{\delta_0}(Z^{\varepsilon_n} \text{ does not hit } R_n A^c)^{[\varepsilon_n^{-1}]} = \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}(X^{\varepsilon_n} \text{ hits } b_n A^c) \tag{3.32}$$

By Theorem 3.3, we know that the law of X^{ε_n} under $\mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}$ converges as $n \to \infty$ to the law of super-Brownian motion with branching mechanism $\psi_{(\kappa)}$ started from δ_0. The next lemma is essentially a consequence of this convergence. We use the notation of Subsection 2.4.

LEMMA 3.13. *We have*

$$\lim_{n\to\infty} \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}(X^{\varepsilon_n} \text{ hits } b_n A^c) = P_{\delta_0}(Y^{(\kappa)} \text{ hits } bA^c),$$

where $b = \sqrt{a} = \lim b_n$.

We postpone the proof of Lemma 3.13 and proceed to the proof of (3.31), in the case when $a > 0$. By the results recalled in Subsection 2.4, we know that

$$P_{\delta_0}(Y^{(\kappa)} \text{ hits } bA^c) = 1 - \exp(-v(0)) \ , \tag{3.33}$$

where the function $(v(x), x \in bA)$ is the unique nonnegative solution of the singular boundary value problem

$$\begin{cases} \frac{1}{2}\Delta v = \psi_{(\kappa)}(u) & \text{in } bA \ , \\ u_{|\partial(bA)} = +\infty \ . \end{cases}$$

It is immediate to verify that $u_{(\kappa a)}(x) = a\,v(bx)$ for every $x \in A$, and in particular $u_{(\kappa a)}(0) = a\,v(0)$.

From (3.32), (3.33) and Lemma 3.13, we obtain

$$\lim_{n\to\infty} (1 - \mathbb{P}_{\delta_0}(Z^{\varepsilon_n} \text{ hits } R_n A^c))^{[\varepsilon_n^{-1}]} = \exp(-v(0))$$

and thus

$$\lim_{n\to\infty} \varepsilon_n^{-1}\,\mathbb{P}_{\delta_0}(Z^{\varepsilon_n} \text{ hits } R_n A^c) = v(0),$$

or equivalently, since $\varepsilon_n R_n^2 \to a$,

$$\lim_{n\to\infty} R_n^2\,\mathbb{P}_{\delta_0}(Z^{\varepsilon_n} \text{ hits } R_n A^c) = av(0) = u_{(\kappa a)}(0).$$

This completes the proof of (3.31), in the case $a > 0$. □

Proof of Lemma 3.13. By replacing A with bA, we may and shall assume in this proof that $b = 1$. We thus have $b_n \to 1$ as $n \to \infty$. We first prove that

$$\liminf_{n\to\infty} \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}(X^{\varepsilon_n} \text{ hits } b_n A^c) \geq P_{\delta_0}(Y^{(\kappa)} \text{ hits } A^c). \tag{3.34}$$

By Lemma 3.8, the events $\{Y^{(\kappa)} \text{ hits } A^c\}$ and $\{Y^{(\kappa)} \text{ hits } \bar{A}^c\}$ coincide a.s. We can then find a countable collection $(\varphi_i)_{i\geq 1}$ of continuous functions with compact support contained in $(\bar{A})^c$, such that

$$\{Y^{(\kappa)} \text{ hits } (\bar{A})^c\} = \Big\{\sup_{i\geq 1}\Big(\sup_{t>0}\langle Y_t^{(\kappa)}, \varphi_i\rangle\Big) > 0\Big\}, \qquad P_{\delta_0} \text{ a.s..}$$

Hence, if $(t_j)_{j\geq 1}$ is a sequence dense in $[0,\infty)$, we have

$$P_{\delta_0}(Y^{(\kappa)} \text{ hits } (\bar{A})^c) = \lim_{N\to\infty} \uparrow P_{\delta_0}\Big(\sup_{1\leq i\leq N}\Big(\sup_{1\leq j\leq N}\langle Y_{t_j}^{(\kappa)}, \varphi_i\rangle\Big) > 0\Big). \tag{3.35}$$

However, Theorem 3.3 implies that, for every $N \geq 1$,

$$\liminf_{n\to\infty}\ \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}\Big(\sup_{1\leq i\leq N}\Big(\sup_{1\leq j\leq N}\langle X_{t_j}^{\varepsilon_n}, \varphi_i\rangle\Big) > 0\Big) \geq P_{\delta_0}\Big(\sup_{1\leq i\leq N}\Big(\sup_{1\leq j\leq N}\langle Y_{t_j}^{(\kappa)}, \varphi_i\rangle\Big) > 0\Big). \tag{3.36}$$

Recall that $b_n \to 1$, and note that the support of each function φ_i is at a strictly positive distance of the set A. As a consequence, for every fixed N, the support of φ_i will be contained in $b_n\bar{A}^c$ for every $i = 1, \ldots, N$, as soon as n is large enough. Hence, for all large enough n,

$$\Big\{\sup_{1\leq i\leq N}\Big(\sup_{1\leq j\leq N}\langle X_{t_j}^{\varepsilon_n}, \varphi_i\rangle\Big) > 0\Big\} \subset \{X^{\varepsilon_n} \text{ hits } b_n A^c\}.$$

Using this inclusion and then (3.36) and (3.35), we immediately obtain (3.34).

We next turn to the more difficult upper bound

$$\limsup_{n\to\infty}\ \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}(X^{\varepsilon_n} \text{ hits } b_n A^c) \leq P_{\delta_0}(Y^{(\kappa)} \text{ hits } A^c). \tag{3.37}$$

We fix $\delta > 0$ small enough so that the closed ball of radius 4δ centered at 0 is contained in A. As in Lemma 3.7, we let A_δ be the connected component of the open set

$$\{x \in A : \operatorname{dist}(x, A^c) > \delta\}$$

which contains 0. We denote the exit measure from A_δ for the rescaled branching Brownian motion X^{ε_n} by $\mathcal{E}_\delta^n$. In other words, the measure $\mathcal{E}_\delta^n$ is equal to ε_n times the sum of the Dirac point masses at all points of ∂A_δ which are first exit points from A_δ for one of the historical paths associated with X^{ε_n} (see Subsection 2.3 for a precise definition of these historical paths).

Let Φ be a continuous function on $\mathbb{R}^d$ such that $0 \le \Phi \le 1$, $\Phi = 0$ on $A_{3\delta}$ and $\Phi = 1$ on $A^c_{2\delta}$. Then, for every $\eta > 0$ and $\rho > 0$,

$$\begin{aligned}\mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}(X^{\varepsilon_n} \text{ hits } b_nA^c) = {}&\mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}(X^{\varepsilon_n} \text{ hits } b_nA^c,\ \langle \mathcal{E}^n_\delta, 1\rangle < \eta)\\ &+ \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}\Big(X^{\varepsilon_n} \text{ hits } b_nA^c,\ \langle \mathcal{E}^n_\delta, 1\rangle \ge \eta,\ \int_0^\infty \langle X^{\varepsilon_n}_s, \Phi\rangle ds \le \rho\Big)\\ &+ \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}\Big(X^{\varepsilon_n} \text{ hits } b_nA^c,\ \langle \mathcal{E}^n_\delta, 1\rangle \ge \eta,\ \int_0^\infty \langle X^{\varepsilon_n}_s, \Phi\rangle ds > \rho\Big). \end{aligned} \tag{3.38}$$

Let $\alpha_n(\eta)$, $\beta_n(\eta,\rho)$ and $\gamma_n(\eta,\rho)$ be the three terms appearing in the right-hand side of (3.38) in this order.

We first bound $\alpha_n(\eta)$. Provided n is sufficiently large, b_nA^c is contained in $A^c_{\delta/2}$ and thus

$$\alpha_n(\eta) \le \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}(X^{\varepsilon_n} \text{ hits } A^c_{\delta/2},\ \langle \mathcal{E}^n_\delta, 1\rangle < \eta).$$

Note that the times at which the historical paths of X^{ε_n} exit A_δ form a stopping line in the sense of [**Ch91**]. We can thus apply the strong Markov property at a stopping line (Proposition 2.1 in [**Ch91**]) to see that $\alpha_n(\delta)$ is bounded above by the probability for a branching Brownian motion (without killing) starting initially with less than $\eta\varepsilon_n^{-1}$ particles, that one of the historical paths reaches a distance greater than $\delta/(2\sqrt{\varepsilon_n})$ from its starting point (to be precise we need a slight extension of the results in [**Ch91**], since our spatial motion is not standard Brownian motion, but Brownian motion killed inside Γ_ϖ). The estimate (3.5) now gives

$$\alpha_n(\eta) \le C_1''(d,\nu)\,\frac{4\eta}{\delta^2}. \tag{3.39}$$

Then, we have

$$\beta_n(\eta,\rho) \le \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}\Big(\langle \mathcal{E}^n_\delta, 1\rangle \ge \eta,\ \int_0^\infty \langle X^{\varepsilon_n}_s, \Phi\rangle ds \le \rho\Big).$$

Recall that $\Phi = 1$ on $A^c_{2\delta}$ and in particular $\Phi = 1$ on $\overline{B}(x,\delta)$ for every $x \in \partial A_\delta$. We use the strong Markov property at the same stopping line as in the previous argument, together with a simple coupling argument, to write that

$$\beta_n(\eta,\rho) \le \mathbb{P}_{[\eta\varepsilon_n^{-1}]\delta_0}\Big(\int_0^\infty \langle \widetilde{X}^{\varepsilon_n}_s, \mathbb{I}_{\overline{B}(0,\delta)}\rangle ds \le \rho\Big),$$

where $\widetilde{X}^{\varepsilon_n}$ is defined in terms of a branching Brownian motion $\widetilde{Z}^{\varepsilon_n}$ in the same way as X^{ε_n} was defined from Z^{ε_n}. This branching Brownian motion $\widetilde{Z}^{\varepsilon_n}$ has the same offspring distribution as Z^{ε_n}, but particles are now killed at rate ε_n homogeneously over $\mathbb{R}^d$. Furthermore, $\widetilde{Z}^{\varepsilon_n}$ also starts from $k\delta_0$ under the probability measure $\mathbb{P}_{k\delta_0}$.

By Proposition 3.4, the law of $(\widetilde{X}^{\varepsilon_n}_t)_{t\ge 0}$ under $\mathbb{P}_{[\eta\varepsilon_n^{-1}]\delta_0}$ converges as $n \to \infty$ to the law of $Y^{(1)}$ under $P_{\eta\delta_0}$ (in the notation of Subsection 2.4), and so

$$\limsup_{n\to\infty} \mathbb{P}_{[\eta\varepsilon_n^{-1}]\delta_0}\Big(\int_0^\infty \langle \widetilde{X}^{\varepsilon_n}_s, \mathbb{I}_{\overline{B}(0,\delta)}\rangle ds \le \rho\Big) \le P_{\eta\delta_0}\Big(\int_0^\infty \langle Y^{(1)}_s, \mathbb{I}_{\overline{B}(0,\delta)}\rangle ds \le \rho\Big) =: \beta_\infty(\eta,\rho).$$

The continuity of sample paths of $Y^{(1)}$ ensures that $\beta_\infty(\eta,\rho) \to 0$ as $\rho \to 0$, for every fixed $\eta > 0$.

For the term $\gamma_n(\eta,\rho)$, we simply use the bound

$$\gamma_n(\eta,\rho) \le \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}\Big(\int_0^\infty \langle X^{\varepsilon_n}_s, \Phi\rangle ds > \rho\Big).$$

Theorem 3.3 implies that

$$\limsup_{n\to\infty} \gamma_n(\eta,\rho) \leq P_{\delta_0}\Big(\int_0^\infty \langle Y_s^{(\kappa)},\Phi\rangle ds \geq \rho\Big) \leq P_{\delta_0}(Y^{(\kappa)} \text{ hits } A_{3\delta}^c),$$

since $\Phi = 0$ on $A_{3\delta}$.

To complete the argument, fix $\vartheta > 0$. By Lemma 3.7 (ii), we can choose $\delta > 0$ sufficiently small so that

$$P_{\delta_0}(Y^{(\kappa)} \text{ hits } A_{3\delta}^c) \leq P_{\delta_0}(Y^{(\kappa)} \text{ hits } A^c) + \frac{\vartheta}{3}.$$

From (3.39), we can then choose $\eta > 0$ sufficiently small so that for all large n,

$$\alpha_n(\eta) \leq \frac{\vartheta}{3}.$$

Finally we choose $\rho > 0$ such that $\beta_\infty(\eta,\rho) \leq \frac{\vartheta}{3}$. From (3.38) and the previous estimates, we obtain

$$\limsup_{n\to\infty} \mathbb{P}_{[\varepsilon_n^{-1}]\delta_0}(X^{\varepsilon_n} \text{ hits } b_n A^c) \leq P_{\delta_0}(Y^{(\kappa)} \text{ hits } A^c) + \vartheta,$$

and since ϑ was arbitrary the proof of (3.37) and of Lemma 3.13 are complete. □

We still have to discuss the case $a = 0$ in (3.31). So, let us consider two sequences $(\varepsilon_n)_{n\geq 1}$ and $(R_n)_{n\geq 1}$ such that $\varepsilon_n R_n^2 \to 0$. Let $a_0 > 0$ and $\varepsilon_n' = \varepsilon_n \vee (a_0 R_n^{-2})$. Since $\varepsilon_n \leq \varepsilon_n'$, we have

$$\liminf_{n\to\infty} R_n^2\, \mathbb{P}_{\delta_0}(Z^{\varepsilon_n} \text{ hits } R_n A^c) \geq \liminf_{n\to\infty} R_n^2\, \mathbb{P}_{\delta_0}(Z^{\varepsilon_n'} \text{ hits } R_n A^c) = u_{(\kappa a_0)}(0),$$

by the case $a > 0$. By Lemma 3.7 (i), $u_{(\kappa a_0)}(0)$ can be made arbitrarily close to $u_{(0)}(0)$ when a_0 is small, and so

$$\liminf_{n\to\infty} R_n^2\, \mathbb{P}_{\delta_0}(Z^{\varepsilon_n} \text{ hits } R_n A^c) \geq u_{(0)}(0).$$

To obtain the corresponding upper bound, a similar coupling argument shows that it suffices to consider the case when $\varepsilon_n = 0$ for every n, that is when there is no killing inside the obstacles. Hence, consider the branching Brownian motion $Z^0 = Z^{\varpi,0}$ (the notation is even more legitimate since $Z^{\varpi,0}$ does not depend on ϖ). For every $\rho > 0$, define a rescaled version of Z^0 by setting

$$\langle \overline{X}_t^{(\rho)}, \varphi\rangle = \rho \int Z^0_{\rho^{-1}t}(dx)\, \varphi(\rho^{1/2}x).$$

By Proposition 3.4, the law of $(\overline{X}_t^{(\rho)})_{t\geq 0}$ under $\mathbb{P}_{[\rho^{-1}]\delta_0}$ converges to the law of $Y^{(0)}$ under P_{δ_0} as ρ tends to 0. Set $\rho_n = R_n^{-2}$, in such a way that

$$\{Z^0 \text{ hits } R_n A^c\} = \{\overline{X}^{(\rho_n)} \text{ hits } A^c\}. \tag{3.40}$$

A simplified version of the arguments of the proof of Lemma 3.13 shows that

$$\limsup_{n\to\infty} \mathbb{P}_{[\rho_n^{-1}]\delta_0}(\overline{X}^{(\rho_n)} \text{ hits } A^c) \leq P_{\delta_0}(Y^{(0)} \text{ hits } A^c) = 1 - \exp(-u_{(0)}(0)).$$

Arguing as in the first part of the proof of the theorem and using (3.40) yields

$$\limsup_{n\to\infty} R_n^2\, \mathbb{P}_{\delta_0}(Z^{\varepsilon_n} \text{ hits } R_n A^c) \leq u_{(0)}(0),$$

which completes the proof of Theorem 3.1. □

Part 2

Collapse of structure in spatial genealogical processes

CHAPTER 4

Genealogies of subdivided populations subject to recurrent mass extinctions

We investigate the infinitely many demes limit of the genealogy of a sample of individuals from a subdivided population that experiences sporadic mass extinction events. By exploiting a separation of time scales that occurs within a class of structured population models generalizing Wright's island model, we show that as the number of demes tends to infinity, the limiting form of the genealogy can be described in terms of the alternation of instantaneous *scattering* phases that depend mainly on local demographic processes, and extended *collecting* phases that are dominated by global processes. When extinction and recolonization events are local, the genealogy is described by Kingman's coalescent, and the scattering phase influences only the overall rate of the process. In contrast, if the demes left vacant by a mass extinction event are recolonized by individuals emerging from a small number of demes, then the limiting genealogy is a coalescent process with simultaneous multiple mergers (a Ξ-coalescent). In this case, the details of the within-deme population dynamics influence not only the overall rate of the coalescent process, but also the statistics of the complex mergers that can occur within sample genealogies. These results suggest that the combined effects of geography and disturbance could play an important role in producing the unusual patterns of genetic variation documented in some marine organisms with high fecundity. These results are a joint work with J.E. Taylor from the University of Oxford, and appeared in [**TV09**].

Contents

1. Introduction

In this chapter, we investigate a class of population genetics models that describe a population of individuals subdivided into D demes which are subject to sporadic mass extinction events.

In general, we will think of these demes as corresponding to geographically distinct subpopulations such as occur in Wright's island model [**Wri31**], but this structure could also arise in other ways, such as through the association of homologous chromosomes within different individuals of a diploid species. Whatever the source of the structure, many species are subject to recurrent disturbances which, if severe enough, can result in the extinction of a large proportion of the population [**Sou84**]. Important sources of widespread disturbance include fire, severe storms, drought, volcanic eruptions, earthquakes, insect outbreaks, and disease epidemics. Our goal in this study is to characterize the effects that such events have on the genealogy of a sample of individuals or genes collected from the entire population. Specifically, we will identify a set of conditions which will guarantee that in the limit of infinitely many demes, the genealogy of the sample converges to a process which alternates between two phases: an extended phase during which ancestral lineages occupy distinct demes, and an effectively instantaneous phase that begins each time two or more lineages are gathered into the same deme and ends when these are again scattered into different subpopulations through a combination of mergers and migrations. The existence of this limit is a consequence of the separation of time scales between demographic events occurring within individual demes and those affecting the global dynamics of the population.

This study was partly motivated by recent investigations of the population genetics of several marine organisms whose genealogies appear to depart significantly from Kingman's coalescent (see Section 1.2). Based on their analysis of sequence polymorphism in a population of the Pacific oyster, [**EW06**] and more recently [**SW08**] suggest that the genealogies of some marine organisms with high fecundity and sweepstakes recruitment may be better described by a class of coalescent processes that generalize Kingman's coalescent by allowing for simultaneous multiple mergers. Indeed, in such organisms, the capacity of individuals to spawn millions of offspring makes it possible, in theory at least, for a substantial fraction of the population to be descended from a single parent. However, depending on the life history and ecology of the species in question, this could happen in several different ways. One possibility is that on rare occasions, individuals give birth to such a large number of offspring that even with random, independent survival of young, these cohorts constitute a sizable proportion of the next generation. Such a scenario has been studied by [**Sch03**], who showed that coalescents with multiple and simultaneous mergers arise naturally when the offspring distribution has a polynomial tail. Another possibility is that on rare occasions a small number of individuals could contribute disproportionately many of the surviving offspring not because they are exceptionally fecund, but because of mass reproductive failure or death in other parts of the population. What distinguishes these two scenarios is whether individuals win the recruitment sweepstakes by producing an exceptionally large number of offspring relative to the long-term average, or by simply giving birth to an average (or even below-average) number of offspring at a time when most other individuals experience an exceptional failure of reproduction.

The multiple mergers that occur in the models investigated in this chapter arise through a combination of both of these factors: mass extinctions create large swathes of unoccupied territory which is then instantaneously repopulated by individuals emerging from finitely many demes. Of course, one weakness of this study is that we do not identify the biological mechanism responsible for restricting recolonization in this way, and in fact it seems difficult to formulate such a mechanism that is both realistic and consistent with the metapopulation models considered in this work. However, there are several scenarios under which similar dynamics could arise in a spatially extended population in which disturbance events tend to affect contiguous demes. For example, if dispersal distances are short, then recolonization of vacant habitat in a one-dimensional population such as along a shoreline or a riparian corridor could be dominated by individuals recently descended from the small number of demes bordering the affected area. Similar reasoning might also apply to organisms with fractal-like distributions, such as aquatic or littoral species in estuarine environments or possibly even HIV-1 populations in the lymphatic system of an infected host. Alternatively, if regrowth from the margins is slow or even impossible (e.g., because surviving demes are separated from vacant demes by inhospitable habitat), then a few long-distance migrants could be responsible for repopulating empty demes even in species with two-dimensional distributions. Furthermore, in this case, we might also predict that the number of demes contributing recolonizers would be negatively correlated with the fecundity of the organism, since less time would be available for additional migrants to enter the affected region before the first migrant propagule had completely repopulated the region. Although the mathematical analysis is much more challenging than that given here, we shall see in Chapter 5 a spatially-explicit model incorporating these features.

1.1. Wright's island model with mass extinctions. To motivate both the class of models studied in this chapter as well as the separation of time scales phenomenon that leads to the infinitely-many demes limit, let us begin by considering a version of Wright's island model with mass extinctions. Suppose that a population of haploid organisms is subdivided into D demes, each of which contains N individuals. We will assume that individuals reproduce continuously, i.e., generations are overlapping, and that at rate 1 each individual gives birth to a single offspring which settles in that same deme with probability $1-m$ and otherwise migrates to one of the other $D-1$ demes, chosen uniformly at random. In either case, we will assume that the deme size is constant and that a newborn individual immediately replaces one of the existing N members of the deme in which it settles. Notice that if $m=0$, then this model reduces to a collection of D independent Moran models in populations of constant size N, whereas if $N=m=1$, it describes the usual Moran model in a single population of size D. However, in the following discussion we will assume that $m>0$ and that D is very much larger than N.

Before we account for mass extinctions, let us consider the genealogy of a sample of n individuals chosen uniformly at random from the entire population. We first observe that, looking backwards in time, each lineage migrates out of its current deme at rate $(D-1)Nm/ND \approx m$. Furthermore, if two lineages occupy different demes, then for these to coalesce, one of the two must migrate into the deme where the other lineage currently resides, an event that occurs approximately at rate m/D; here we have neglected terms of order D^{-2} and will continue to do so without further comment. When two lineages are collected in the same deme, then they can either coalesce immediately, which happens with probability $1/N$, or they can cohabit within that deme for some random period of time until either they coalesce or they migrate into different demes. Since two lineages occupying the same deme coalesce at rate $2(1-m)/N$, and each lineage, independently of the other, migrates out of the deme at rate m, the probability that the two coalesce rather than migrate is $\chi = (1-m)/(1-m+Nm)$. Putting these observations together, it follows that every time two lineages are collected within the same deme by migration, the total probability that they coalesce rather than migrate into different demes is $1/N + (1-1/N)\chi = 1/(1-m+Nm)$, and the time that elapses between entry into the same deme and either coalescence or escape is a mixture of a point mass at 0 (in case they coalesce at the entry time) and an exponential random variable with mean $N/(2mN + 2(1-m))$. In particular, notice that typically much less time is required for two lineages occupying the same deme to either coalesce or escape (of order N) than is needed for two lineages occupying different demes to be collected into the same deme (of order D). It is this disparity between the rate of events happening within individual demes and the rate at which lineages are gathered together that gives rise to a separation of time scales in the island model. If we rescale time by a factor of D and let the number of demes tend to infinity, then the time required for two lineages sampled from different demes to coalesce is exponentially distributed with mean $(1 - m + Nm)/2m$.

To complete our description of the coalescent process in this model, we need to consider the possibility of more complex coalescent events. We first observe that if n individuals are sampled from D demes, then the probability that all of these individuals reside in different demes will be close to one if D is much greater than n. Furthermore, because lineages occupying different demes coalesce and migrate independently of one another, it is straightforward to show that the probability that three or more lineages ancestral to our sample are collected into the same deme is of order D^{-2} or smaller. Likewise, it can be shown that the probability of having multiple pairs of lineages collected into several demes at the same time is similarly negligible. From these observations, it follows that only pairwise coalescence events matter in the infinitely-many demes limit, and that if there are n ancestral lineages, then at rate $\binom{n}{2}2m/(1 - m + Nm)$, two of these, chosen uniformly at random, coalesce, leaving $n - 1$ ancestral lineages. In other words, the genealogy for this model can be approximated by a scalar time change of Kingman's coalescent, with a rate that depends on both the migration rate and the deme size. This result is essentially due to [**WA01**], who considered a similar model with non-overlapping generations and Wright-Fisher sampling.

Now let us introduce mass extinction events into this model. Fix $e > 0$ and $y \in [0, 1)$, and suppose that at rate e/D, the metapopulation suffers a disturbance which causes each deme to go extinct, independently of all others, with probability y. For example, we could consider a model in which the demes represent small islands or keys in the Caribbean and the disturbances are hurricanes that completely inundate those islands lying in their path. Here we are reverting to the original time units, i.e., time has not yet been rescaled by a factor of D, and we have chosen the disturbance rate so that mass extinctions occur at rates commensurate with coalescence in the pure island model. In keeping with the assumption that deme size is constant, we will assume that all of the islands that are left vacant by a mass extinction are immediately recolonized by offspring dispersing out of a single source deme that is chosen uniformly at random from among the demes unaffected by the disturbance. In addition, we will assume that the parent of each colonizing individual is chosen uniformly at random from the N members of the source deme. Of course, the entire metapopulation could be extirpated by a mass extinction if $y > 0$, but the probability of this outcome is exponentially small in D and can be disregarded as D tends to infinity.

Suppose that a mass extinction occurs at a time when there are n ancestral lineages occupying distinct demes. Bearing in mind that we are now looking backwards in time, all of the lineages belonging to demes that are affected by the disturbance will move into the source deme, where those sharing the same parent will immediately coalesce. Thus, one reason that multiple mergers can occur in this model is because of the very highly skewed distribution of recolonizing offspring contributed both by individuals and demes following a mass extinction. Suppose that there are n_1 distinct lineages remaining in the source deme once we account for this initial set of coalescences. These lineages will undergo a random sequence of migration and coalescence events until there is only one lineage remaining within the source deme. For example, if $n_1 = 4$, then one possible outcome would see one lineage migrate out of the deme followed by a pair of binary mergers, leaving only one lineage within the source deme. Whatever the sequence, the amount of time required to scatter the lineages into different demes will be of order $\mathcal{O}(1)$, whereas the time until either the next mass extinction event or the next binary merger involving lineages outside of the source deme will be of order $\mathcal{O}(D)$. Thus, if we again rescale time by a factor of D, then any sequence of coalescence and migration events involving a source deme will effectively be instantaneous when we let D tend to infinity. This is the second way in which multiple merger events can arise in this model. Furthermore, varying the migration rate and deme size changes not only the overall rate of coalescence, but also the relative rates of the different kinds of multiple merger events that can occur. For example, if Nm is very small, then the coalescent process will be close to a Λ-coalescent (which has multiple mergers, but not simultaneous multiple mergers) because most lineages that are collected into a source deme by a mass extinction event will coalesce before any escape by migration. However, as N increases, so will the probability that multiple lineages enter into and then escape from the source deme without coalescing. This suggests that at moderate values of Nm, mass extinctions may be likely to result in simultaneous mergers (i.e., the coalescent is a Ξ-coalescent), while for very large values of Nm, multiple mergers of all types will be unlikely and the coalescent process will tend towards Kingman's coalescent.

1.2. Neutral genealogies and coalescents. In the last twenty years, coalescent processes have taken on increasingly important role in both theoretical and applied population genetics, where their relationship to genealogical trees has made them powerful tools to study the evolution of genetic diversity within a population. Under the assumption of neutrality, allelic types do not influence the reproduction of individuals and it is therefore possible to separate 'type' and 'descent'. This allows us to study the genealogy of a sample of individuals on its own and then superimpose a mechanism describing how types are transmitted from parent to offspring, justifying the interest in investigating genealogical processes corresponding to particular reproduction mechanisms without explicit mention of types. We refer to [**Nor01**] for a review of coalescent theory in population genetics.

Beginning with the coalescent process introduced by [**Kin82**] to model the genealogy of a sample of individuals from a large population, three increasingly general classes of coalescent processes have been described. A key feature shared by all three classes is the following consistency property: the process induced on the set of all partitions of $\{1,\dots,n\}$ by the coalescent acting on the partitions of $\{1,\dots,n+k\}$ (obtained by considering only the blocks containing elements of $\{1,\dots,n\}$) has the same law as the coalescent acting on the partitions of $\{1,\dots,n\}$. In terms of genealogies, this property means that the genealogy of n individuals does not depend on the size of the sample that contains them. To describe these continuous-time Markov processes, it will be convenient to introduce some notation. For all $n \in \mathbb{N}$, we denote the set of all partitions of $[n] := \{1,\dots,n\}$ by $\mathcal{P}_n$. In the following, the index n of the set of partitions in which we are working will be referred to as the *sample size*, an element of $\{1,\dots,n\}$ will be called an *individual*, and 'block' or '*lineage*' will be equivalent terminology to refer to an equivalence class. If $\zeta \in \bigcup_n \mathcal{P}_n$, then $|\zeta| = k$ means that the partition ζ has k blocks. Also, for $\zeta, \eta \in \mathrm{P}_n$ and $k_1,\dots,k_r \geq 2$, we will write $\eta \subset_{k_1,\dots,k_r} \zeta$ if η is obtained from ζ by merging exactly k_1 blocks of ζ into one block, k_2 into another block, and so on. Kingman's coalescent is defined on $\mathcal{P}_n$ for all $n \geq 1$, as a Markov process with the following Q-matrix: if $\zeta, \eta \in \mathcal{P}_n$,

$$q_K(\zeta \to \eta) = \begin{cases} 1 & \text{if } \eta \subset_2 \zeta, \\ -\binom{|\zeta|}{2} & \text{if } \eta = \zeta, \\ 0 & \text{otherwise.} \end{cases}$$

A more general class of exchangeable coalescents, allowing mergers of more than two blocks at a time, was studied by [**Pit99**] and [**Sag99**]. These coalescents with multiple mergers (or Λ-coalescents) are in one-to-one correspondence with the finite measures on $[0,1]$ in the following manner: for a given coalescent, there exists a unique finite measure Λ on $[0,1]$ such that the entries

$q_\Lambda(\zeta \to \eta)$ of the Q-matrix of the coalescent, for $\zeta, \eta \in \mathcal{P}_n$, are given by

$$q_\Lambda(\zeta \to \eta) = \begin{cases} \int_0^1 \Lambda(dx) x^{k-2}(1-x)^{b-k} & \text{if } \eta \subset_k \zeta \text{ and } |\zeta| = b, \\ -\int_0^1 \Lambda(dx) x^{-2}\big(1-(1-x)^{b-1}(1-x+bx)\big) & \text{if } \eta = \zeta \text{ and } |\zeta| = b, \\ 0 & \text{otherwise.} \end{cases}$$

Kingman's coalescent is recovered by taking $\Lambda = \delta_0$, the point mass at 0. Lastly, a third and wider class of coalescents was introduced by [**MS01**] and [**Sch00**], for which mergers involving more than one ancestor are allowed. These coalescents with simultaneous multiple mergers (or Ξ-coalescents) are characterized in [**Sch00**] by a finite Borel measure on the infinite ordered simplex

$$\Delta = \Big\{(x_1, x_2, \ldots) : x_1 \geq x_2 \geq \ldots \geq 0, \sum_{i=1}^{\infty} x_i \leq 1\Big\}.$$

Indeed, to each coalescent corresponds a unique finite measure Ξ on Δ of the form $\Xi = \Xi_0 + a\delta_{\mathbf{0}}$, where Ξ_0 has no atom at zero and $a \in [0, \infty)$, such that the transition rates of the coalescent acting on $\mathcal{P}_n$ are given by

$$q_\Xi(\zeta \to \eta) = \int_\Delta \frac{\Xi_0(d\mathbf{x})}{\sum_{j=1}^{\infty} x_j^2} \Bigg(\sum_{l=0}^{s} \sum_{i_1 \neq \ldots \neq i_{r+l}} \binom{s}{l} x_{i_1}^{k_1} \ldots x_{i_r}^{k_r} x_{i_{r+1}} \ldots x_{i_{r+l}} \Big(1 - \sum_{j=1}^{\infty} x_j\Big)^{s-l} \Bigg)$$
$$+a\, \mathbb{I}_{\{r=1, k_1=2\}}$$

if $\eta \subset_{k_1,\ldots,k_r} \zeta$ and $s := |\zeta| - \sum_{i=1}^{r} k_i$. The other rates (for $\eta \neq \zeta$) are equal to zero. The Λ-coalescents are particular cases of Ξ-coalescents, for which $\Xi(x_2 > 0) = 0$.

As mentioned above, coalescent processes can be used to describe the genealogy of large populations. Indeed, a large body of literature has been devoted to describing conditions on the demography of a population of finite size N that guarantee that the genealogical process of a sample of individuals converges to a coalescent as N tends to infinity. Such limiting results for populations with discrete non-overlapping generations are reviewed in [**Möh00**], and some examples can be found for instance in [**Sch03**], [**EW06**] and [**SW08**]. In these examples, the shape of the limiting coalescent is related to the propensity of individuals to produce a non-negligible fraction of the population in the next generation.

However, the representation of the genealogy as a coalescent requires in particular that any pair of lineages has the same chance to coalesce. This condition breaks down when the population is structured into subpopulations, since then coalescence will occur disproportionately often between lineages belonging to the same deme. To model these kinds of scenarios, structured analogues of coalescent processes were introduced (see e.g. [**Not90, WH98**]), which allow lineages both to move between demes as well as coalesce within demes. Various state spaces have been used to describe a structured coalescent, such as vectors in which the i'th component gives the lineages (or their number) present in deme i, or vectors of pairs 'block × deme label'. All these representations of a structured genealogy take into account the fact that the reproductive or dispersal dynamics may differ between demes, hence the need to keep track of the location of the lineages. In contrast, several papers investigate models where the structure of the genealogy collapses on an appropriate time scale, i.e., the limiting genealogy no longer sees the geographical division of the population. In [**Cox89**], demes are located at the sites of the torus $\mathbb{T}(D) \subset \mathbb{Z}^d$ of size D and each site can contain at most one lineage. Lineages move between sites according to a simple random walk, and when one of them lands on a site already occupied, it merges instantaneously with the inhabitant of this 'deme'. These coalescing random walks, dual to the voter model on the torus, are proved to converge to Kingman's coalescent as $D \to \infty$. More precisely, Cox shows that if $n < \infty$ lineages start from n sites independently and uniformly distributed over $\mathbb{T}(D)$, then the process counting the number of distinct lineages converges to the pure death process that describes the number of lineages in Kingman's coalescent. This analysis is generalized in [**CD02**] and [**ZCD05**], where each site of the torus now contains $N \in \mathbb{N}$ individuals and a Moran-type reproduction dynamics occurs within each deme. Again, the limiting genealogy of a finite number of particles sampled at distant sites is given by Kingman's coalescent, and convergence is in the same sense as for Cox' result. Other studies of systems of particles moving between discrete subpopulations and coalescing do not require that the initial locations of the lineages be thinned out. In [**GLW07**], demes are distributed over the grid $\mathbb{Z}^2$ and the process starts with a Poisson-distributed number of lineages on each site of a large box of size $D^{\alpha/2}$, for some $\alpha \in (0,1]$. The authors show that the total number of lineages alive at times of the form D^t converges in distribution as a process (indexed by $t \geq \alpha$) to a time-change of the block counting process of Kingman's coalescent, as $D \to \infty$. See [**GLW07**] for many other references related to these ideas.

Our emphasis in this work will be on the separation of time scales phenomenon and the way in which local and global demographic processes jointly determine the statistics of the limiting coalescent process. Consequently, we shall always assume that the demes comprising our population are exchangeable, i.e., the same demographic processes operate within each deme, and migrants are equally likely to come from any one of the D demes. In this simplified setting, we only need to know how lineages are grouped into demes, but not the labels of these demes.

1.3. Separation of time scales.

A separation of time scales can be said to occur whenever different components of a stochastic process evolve at rates which greatly differ in their magnitudes. This concept is usually invoked when there is a sequence of stochastic processes $(X_t^D, t \geq 0)$ on a space E as well as a function $\eta : E \to E'$ and an increasing sequence $r_D \to \infty$ such that the processes $(X^D_{r_D^{-1}t}, t \geq 0)$ have a non-trivial limit $(X_t^{\text{fast}}, t \geq 0)$ determined by the *fast* time scale, while the processes $(\eta(X_t^D), t \geq 0)$ (which are only weakly influenced by the fast evolution) have another non-trivial limit $(X_t^{\text{slow}}, t \geq 0)$ determined by the *slow* time scale. It is often the case that the processes $(X_t^D, t \geq 0)$ have the Markov property but do not converge to a limit, while the slow processes $(\eta(X_t^D), t \geq 0)$ do converge, but are not Markovian.

Separation of time scales techniques were first introduced into population genetics by [**EN80**], and since then have been used to study the genealogical processes of structured populations in several different settings. In [**NK02**], Nordborg and Krone consider a population of total size N, evolving according to a Wright-Fisher model (see [**Fi30, Wri31**]) and distributed over $D < \infty$ demes. These demes are in turn structured into groups of demes, within which individuals migrate faster (at a rate of order $N^{-\alpha}$ for an $\alpha \in [0,1]$) than from one group to another (which occurs at a rate $\mathcal{O}(N^{-1})$). When all demes are connected by fast migration, they show that structured genealogy collapses to an unstructured Kingman's coalescent as N tends to infinity, due to the fact that migration is so fast compared to the coalescence rate (of order N^{-1}) that the population becomes well-mixed before the first coalescence event occurs. When several groups of demes are connected by slow migration, the genealogical process converges to a structured coalescent, in which groups of demes act as panmictic populations and coalescence of lineages within a group is faster than between two groups. These results are made possible by the fact that the blocks of the partition induced by the genealogy are not affected by a migration event. Since only migration occurs on the fast time scale and coalescence is on the slow time scale, forgetting about the location of the lineages gives a sequence of (non-Markov) processes which converge on the slow time scale to a Markov process.

Another kind of separation of time scales was studied by Wakeley and co-authors in a series of papers (see in particular [**Wak98, Wak99, Wak04, WA01**]). In these models, a population evolving in discrete non-overlapping generations occupies D demes, labeled $1, \dots, D$. Deme i contains a population of N_i adults and receives M_i migrants each generation. Then, a Wright-Fisher resampling within each deme brings the population sizes back to their initial values. Other mechanisms can also be taken into account, such as extinction of a group of demes followed by instantaneous recolonization. Allowing D to tend to infinity greatly simplifies the analysis of the genealogical processes, and in particular gives rise to a decomposition of the genealogy of a sample of individuals into two different phases, occurring on two time scales. Following the terminology introduced in [**Wak99**], the first phase to occur is the *scattering phase*, in which lineages occupying the same deme coalesce or move to *empty* demes ('empty' meaning that none of the sampled lineages are in this deme). In the limit, this phase occurs on the fast time scale and is therefore viewed as instantaneous. At the end of the scattering phase, all remaining lineages lie in different demes. The *collecting phase* is the following period of time during which lineages are gathered together into the same demes by migration or extinction/recolonization, where they may merge. The limiting genealogical process is a coalescent on the slow time scale, which ends when the number of lineages reaches one.

As we have already mentioned, apart from an initial instantaneous burst of mergers (which only occurs if multiple individuals are sampled from the same deme) all of the genealogical processes obtained in this setting are scalar time changes of Kingman's coalescent. Indeed, in the forwards in time evolution, migrants and colonizers are assumed to come from the whole population or from a non-vanishing fraction of the demes and so, with probability one, only two of the finitely many lineages of the sample are brought into the same deme at a time in the limit. Subsequently, the two lineages either coalesce or are scattered again, but in any case the outcome is at most a binary merger. In this chapter, we shall study coalescent processes that arise in population models which include mass extinctions and general recolonization mechanisms, and describe the conditions in which it corresponds to an unstructured Ξ-coalescent on the slow time scale. To this end, we will speak of 'scattering' and 'collecting' phases in a more general sense. We prefer to call the 'collecting phase' the period of time during which lineages wander among empty demes until a migration or extinction event brings several lineages into the same deme. We shall show that, once such a 'geographical collision' has occurred, an instantaneous scattering phase follows at the end of which all lineages have merged or moved to empty demes. Another collecting phase then starts and so on until the most recent common ancestor of the sample has been reached and there is only one lineage remaining.

1.4. Framework and main results. Fix $n \in \mathbb{N}$ and consider the genealogy of a sample of n individuals from a population of $D > n$ demes (the following framework also allows $D = \infty$). In the following, we shall suppose that demes are exchangeable in the sense given in Section 1.2. We shall work in the space P_n^s defined as follows:

DEFINITION 4.1. *Let* $\tilde{\mathrm{P}}_n^s$ *be the set*

$$\tilde{\mathrm{P}}_n^s := \big\{ \big(\{B_1, \dots, B_{i_1}\}, \dots, \{B_{i_{n-1}+1}, \dots, B_{i_n}\}\big) \, : \, 0 \le i_1 \le \dots \le i_n \le n, \\ \emptyset \ne B_j \subset [n] \; \forall j \in \{1, \dots, i_n\}, \; \{B_1, \dots, B_{i_n}\} \in \mathcal{P}_n \big\}$$

of n*-tuples of sets (we allow some of the components of the* n*-tuple to be empty), and let us define the equivalence relation* $\sim$ *on* $\tilde{\mathrm{P}}_n^s$ *by* $\xi \sim \xi'$ *if and only if there exists a permutation* σ *of* $[n]$ *such that, if* $\mathcal{B}_1 := \{B_1, \dots, B_{i_1}\}, \dots, \mathcal{B}_n := \{B_{i_{n-1}+1}, \dots, B_{i_n}\}$ *are the components of the vector* ξ*, then* $\xi' = (\mathcal{B}_{\sigma(1)}, \dots, \mathcal{B}_{\sigma(n)})$*. The quotient of* $\tilde{\mathrm{P}}_n^s$ *by* $\sim$ *is denoted by* P_n^s*.*

We call any $(\{B_1,\dots,B_{i_1}\},\dots,\{B_{i_{n-1}+1},\dots,B_{i_n}\})\in \mathrm{P}^s_n$ *an* unordered structured partition of $[n]$.

In view of the application we have in mind, each component $\mathcal{B}_j$ represents a particular deme containing some of the lineages ancestral to the sample, and the blocks B_k (for $k\in\{1,\dots,i_n\}$) specify the partition of the sample determined by the ancestors alive at a particular time. Empty components are used to guarantee a constant vector size, n, independent of the index D used later. In the following, we omit the term 'unordered' when referring to the structured partitions of Definition 4.1.

The finite set P^s_n is endowed with the discrete topology, which is equivalent to the quotient by $\sim$ of the discrete topology on $\tilde{\mathrm{P}}^s_n$.

DEFINITION 4.2. *A Markov process* $\mathcal{A}$ *on* P^s_n *for which blocks can only merge and change component is called a* structured genealogical process.

To illustrate the possible transitions, let us take $n=5$ and consider the following sequence of events:

$$\begin{aligned}(\{\{1\}\},\{\{2\}\},\{\{3\}\},\{\{4\}\},\{\{5\}\}) &\overset{(i)}{\to} (\{\{1\},\{2\}\},\{\{3,4\}\},\{\{5\}\},\emptyset,\emptyset)\\ &\overset{(ii)}{\to} (\{\{1\}\},\{\{2\}\},\{\{3,4,5\}\},\emptyset,\emptyset)\\ &\overset{(iii)}{\to} \big(\{\{1,2,3,4,5\}\},\emptyset,\emptyset,\emptyset,\emptyset\big).\end{aligned}$$

In this example, we start from the configuration in P^s_5 where each lineage is alone in its deme. During transition (i), either $\{1\}$ or $\{2\}$ changes component and both blocks end up in the same deme (which creates an empty component in our representation), but remain distinct. In contrast, either $\{3\}$ or $\{4\}$ also moves (emptying another component), but then the two blocks merge into a single block $\{3,4\}$ which is not allowed to split during later transitions. Block $\{5\}$ remains alone in its component. During transition (ii), lineages $\{1\}$ and $\{2\}$ are scattered again into two different demes by the movement of one of them, while one of the lineages $\{3,4\}$ or $\{5\}$ changes component and the two blocks merge. Eventually, all the remaining blocks are gathered into the same deme and merge into a single block. Since elements of P^s_5 are defined up to a permutation of their components and since a block is not allowed to split, no other change is possible from the state reached after transition (iii).

REMARK 4.3. *Movements and mergers of blocks do not alter the sample size. However, this does not guarantee that the structured genealogies are consistent in the sense given in Section 1.2 as we would expect from a reasonable genealogical process. In fact, several conditions will be imposed on the models we consider so that this property holds: see Lemma 4.12 for the consistency of the fast genealogical process, and the set of conditions (4.4) imposed on the geographical gatherings in Proposition 4.16. Proposition 4.25 states in particular that the latter conditions are necessary and sufficient for the genealogies to be consistent on both time scales and that when they are fulfilled, the unstructured genealogical process on the slow time scale is a* Ξ*-coalescent.*

Let us order the components of a given structured partition by the smallest element belonging to a block contained in the component (if it is non-empty). Empty components come last. For each $k\le n$ and $\zeta\in\mathrm{P}^s_n$, let us write $|\zeta|_a=k$ if the a'th component (in the order just defined) of the structured partition ζ contains k blocks, and define a subset Π_n of P^s_n by

$$\Pi_n := \big\{\zeta\in\mathrm{P}^s_n : |\zeta|_a\le 1 \quad \forall\, a\in\{1,\dots,n\}\big\}. \tag{4.1}$$

Π_n is the set of all structured partitions of $[n]$ in which each deme contains at most one lineage. These sets will appear naturally in the description of the limiting processes.

Recall from the example given in Section 1.1 that the rate at which lineages are collected together in the same deme is much smaller than the rate at which lineages already occupying the same deme either coalesce or are scattered into different demes. Furthermore, as in that example, we will continue to assume that catastrophic extinction-recolonization events occur rarely, in fact, at rates that are of the same order of magnitude as the rate at which lineages occupying different demes are brought together by ordinary migration. With these points in mind, let us consider a sequence $(\mathcal{A}_s^D, s \geq 0)$ of structured genealogical processes for a finite sample from the whole population, which consists of the following kinds of events:

- within-deme coalescence and movement of lineages to empty demes at rates of order $\mathcal{O}(1)$;
- movement of groups of lineages initially occupying different demes into the same deme, possibly followed by mergers of some of these lineages, at rates of order $\mathcal{O}(r_D^{-1})$.

Let us rescale time by a factor of r_D so that the coalescence rate of two individuals in different demes is of order $\mathcal{O}(1)$ as D tends to infinity. Of course, within-deme coalescence and migration now occur at increasing rates of order $\mathcal{O}(r_D)$. This implies that, for a given sample size n, the generator G^D of the genealogical process acting on P_n^s has the form

$$G^D = r_D \Psi + \Gamma + R_D,$$

where Ψ, Γ and R_D are bounded linear operators, $\langle R_D \rangle \to 0$ as $D \to \infty$, and we do not record the dependence of the operators on the sample size n. Here, if $\|\cdot\|$ stands for the supremum norm on the space of functions $f : \mathrm{P}_n^s \to \mathbb{R}$, then $\langle R \rangle$ is defined by

$$\langle R \rangle = \sup_{f \neq 0} \frac{\|Rf\|}{\|f\|}. \tag{4.2}$$

Because $r_D \to \infty$, the sequence $(G^D)_{D\geq 1}$ is unbounded, even when applied to functions of the unstructured partition induced by $\mathcal{A}^D$, and so we do not expect the structured coalescent processes corresponding to these generators to converge pathwise. Nevertheless, our heuristic description of the fast dynamics suggests that elements of Π_n will be unaffected by the 'fast' events corresponding to Ψ, which will indeed be the case under the assumptions made in Section 2. Furthermore, we will show (cf. Lemma 4.10) that the process generated by Ψ on P_n^s and starting at $\zeta \in \mathrm{P}_n^s$ a.s. reaches a random final state $\underline{\zeta}$ in Π_n in a finite number of steps. Since the rates of the events generated by Ψ grow to infinity, increasing numbers of these events take place before the first event corresponding to Γ even occurs. This motivates the description of the genealogy given above in terms of an alternation of very short scattering phases driven by Ψ and of longer collecting phases ending with the first event generated by Γ at which $\mathcal{A}^D$ leaves Π_n. Viewing all of the transitions occurring during a given scattering phase as a single, more complex event, and exploiting the fact that these phases are vanishingly short, it is plausible that there is a genealogical process $\mathcal{A}$ with values in Π_n such that for each fixed time $t > 0$, $\mathcal{A}_t^D \Rightarrow \mathcal{A}_t$ as $D \to \infty$. Our main result makes these heuristic arguments rigorous:

THEOREM 4.4. *Let $\zeta \in \mathrm{P}_n^s$. Under the conditions described in Section 3.1, the finite-dimensional distributions of the structured genealogical process $\mathcal{A}^D$ starting at ζ converge to those of a Π_n-valued Markov process $\mathcal{A}$ starting at $\underline{\zeta}$, except at time* 0.

The proof that $\mathcal{A}^D$ converges in law to $\mathcal{A}$ in the Skorokhod space $D_{\mathrm{P}_n^s}[0,\infty)$ of all càdlàg paths with values in P_n^s requires tightness of the corresponding sequence of distributions. We shall show in Proposition 4.22 that this property holds if and only if the rate at which the genealogical process leaves the set Π_n tends to zero as D grows to infinity. Indeed, if this condition is not satisfied, then two or more jumps can accumulate during a scattering phase: the jump out of Π_n followed by the events needed to bring $\mathcal{A}^D$ back into Π_n. Fortunately, the proof that the unstructured genealogical processes are tight is less demanding, since these processes do not change state when lineages move between demes. In this case, an accumulation of jumps due to the fast within-deme dynamics will be ruled out if we can show that the probability that the process $\mathcal{A}^D$ re-enters Π_n in a single jump converges to one as D tends to infinity.

The limiting process $\mathcal{A}$ with values in Π_n is introduced and investigated in Section 2, and we show in Proposition 4.16 that, under the assumptions of Theorem 4.4, the unstructured genealogical process induced by $\mathcal{A}$ is the restriction to $\mathcal{P}_n$ of a Ξ-coalescent. We also identify the limiting process ξ for the genealogy on the fast time scale in Section 2, and state in Proposition 4.18 the convergence of $\mathcal{A}^D_{r_D^{-1}\cdot}$ to ξ as processes with values in $D_{\mathrm{P}^s_n}[0,\infty)$. The proofs of Theorem 4.4 and Proposition 4.18 are given in Section 3, along with a discussion of the tightness of $\mathcal{A}^D$. Although the conditions of Theorem 4.4 are somewhat contrived, we show in Section 4 that these are necessary and sufficient for the unstructured genealogical process of a generalized island model to converge to a Ξ-coalescent on the slow time scale. In Section 5, we apply these results to a particular class of models incorporating mass extinction events. Based on our analysis of this class, we suggest that families of Ξ-coalescents may often interpolate between Λ-coalescents and Kingman's coalescent in structured population models, and that it may be a generic property of such models that they admit simultaneous mergers whenever they admit multiple mergers.

2. Construction of the limiting genealogical processes

2.1. A generalized Island-Cannings Model. To motivate the genealogical processes considered in this work, we begin by introducing a general model for the demography of a subdivided population which combines features of the Cannings model [**Can74**] with those of the classical island model [**Wri31**].

Suppose that the population is subdivided into D demes, each of which contains N haploid individuals. Islands are labeled $1,\dots,D$, while individuals within each island are labeled $1,\dots,N$. At rate 1, an ND^2-dimensional random vector $R := \big(R^{i,j}_k,\ i,j \in \{1,\dots,D\}, k \in \{1,\dots,N\}\big)$ is chosen, such that for all i,j,k, $R^{i,j}_k$ is the number of descendants of the k'th individual in deme j which settle into deme i during the event. In keeping with the spirit of the Cannings' model, we use the term 'descendant' both to refer to the offspring of reproducing individuals as well as to individuals which were alive both before and after the event (as in Cannings' formulation of the Moran model). We impose the following conditions on the random variables $R^{i,j}_k$:

1. **Constant deme size:** With probability 1, for all $i \in [D]$ we have $\sum_{j,k} R^{i,j}_k = N$.
2. **Exchangeable dynamics:** The law of R is invariant under any permutation σ of $[D]^2 \times [N]$ such that for every $i \in [D]$, $\sigma(i,i,k)_1 = \sigma(i,i,k)_2$, i.e., σ conserves the relation *source deme* $=$ *destination deme*. (Here, $\sigma(i,j,k)_l$ denotes the l'th component of the permuted vector.)

Then, in each deme the current population is replaced by the N offspring coming into this deme during the event, which we label in an exchangeable manner.

Let us comment on the above conditions. The first one simply guarantees that the number of individuals in each deme is constant and equal to N. For the second condition, let us first fix $i,\ j$ and a permutation τ of $[N]$, and look at the permutation σ given by $\sigma(i,j,k) = (i,j,\tau(k))$ and $\sigma(i',j',k') = (i',j',k')$ whenever $i \neq i'$ or $j \neq j'$. Then, Condition 2 corresponds to the exchangeability of the contribution of the inhabitants of deme j in repopulating deme i. Second, fix

i and choose a permutation τ of $[D]\setminus\{i\}$. Set $\sigma(i,j,k)=(i,\tau(j),k)$ if $j\neq i$, $\sigma(i,i,k)=(i,i,k)$ and $\sigma(i',j',k')=(i',j',k')$ whenever $i'\neq i$. In this case, Condition 2 states that the demes different from deme i contribute in an exchangeable manner to the repopulation of deme i. Finally, let τ be a permutation of $[D]$ and define $\sigma(i,j,k)=(\tau(i),j,k)$ if $j\notin\{i,\tau(i)\}$, $\sigma(i,j,k)=(\tau(i),\tau(i),l)$ if $j=i$, and $\sigma(i,j,k)=(\tau(i),i,l)$ if $j=\tau(i)$. For such permutations, Condition 2 asserts that the dispersal mechanism is exchangeable with respect to the destination of dispersing individuals (provided that this differs from the source deme). Overall, our assumptions aim at making the dynamics depend on the labels as weakly as possible, but we allow the repopulation mechanism of a deme to differ according to whether the new individuals are produced within this deme or come from one of the $D-1$ other demes.

EXAMPLE 4.5. *If R is invariant under all permutations σ of $[D]^2\times[N]$ (not just those satisfying Condition 2), then the dynamics are those of a Cannings' model for a panmictic population of size DN, i.e., there is no population subdivision.*

EXAMPLE 4.6. *If all demes evolve independently of each other, then $R^{i,j}\equiv 0$ whenever $j\neq i$. Condition 2 imposes that $(R^{i,i}, i\in[D])$ should be an exchangeable D-tuple of exchangeable N-tuples, a situation corresponding to a continuous-time Cannings model acting within each deme.*

EXAMPLE 4.7. *Let $m\in[0,1]$ and assume that, with probability $1-m$, R is chosen as in Example 4.6. With probability m, four numbers i,j,l,k are sampled uniformly at random in $[D]^2\times[N]^2$, and the k'th individual in deme j produces an offspring that replaces the l'th individual in deme i. In this case, $R^{i,i}=(1,\dots,0,\dots,1)$, where the unique zero is in the l'th coordinate; $R^{i,j}=(0,\dots,1,\dots,0)$, where the unique 1 is in the k'th coordinate; $R^{i,j'}=(0,\dots,0)$ if $j'\notin\{i,j\}$ and for $i'\neq i$, $R^{i',j'}=(1,\dots,1)$ if $i'=j'$ and $(0,\dots,0)$ otherwise. This model gives a simple example including within-deme reproduction and individual migration. Alternatively, individuals could be exchanged between demes during a migration event, in which case a descendant of individual l in deme i (in the above notation) also replaces individual k in deme j.*

EXAMPLE 4.8. *An event during which one deme goes extinct and is recolonized by the offspring of individuals coming from other demes has the following formulation: $R^{i,i}=(0,\dots,0)$ if deme i goes extinct, $R^{l,l}=(1,\dots,1)$ if $l\neq i$ and the repopulation of deme i satisfies the exchangeability Condition 2. For instance, N individuals are chosen uniformly at random among the $N(D-1)$ inhabitants of the other demes and contribute one offspring in the new population of deme i.*

Many other kinds of events can be imagined, but these three mechanisms (reproduction, migration and extinction/recolonization) will be the building blocks of the models we shall consider in this chapter. Viewed backwards in time, reproduction events as in Example 4.6 will correspond to the merger of several lineages if they are produced (forwards in time) by the same individual during the event considered. A migration event such as in Example 4.7 will correspond to the movement of one or a few lineages from their demes to other subpopulations, if these lineages happen to have their parents in the source demes. An extinction event will also typically result in the movement of lineages among demes, and could involve much larger numbers of individuals or demes than simple migration events. Note that lineages can both move and merge during the same event, if their common parent lies in a different deme.

2.2. Genealogy on the fast time scale. Let us start by constructing a structured genealogical process $(\xi_t, t\geq 0)$ such that its restriction to P_n^s describes the genealogy of n individuals on the fast time scale of individual demes. This process will incorporate mergers of lineages occupying the same deme as well as dispersal of lineages into empty demes (i.e., those not containing other ancestral lineages), but no events where geographically separated lineages end up in identical demes and possibly merge. In fact, if the rate at which such events occur is very large, then it is not difficult to see that the structure of the population effectively disappears on the fast time scale and the model collapses to that of a panmictic population. We thus rule out this kind of situations to keep a structured population.

We construct the process ξ by specifying its restriction to P_n^s. As P_n^s is a finite set, we can define a continuous-time Markov process on this space by specifying its transition rates. Because a block represents a single ancestor, whose descendance at time 0 is made of the individuals contained in the block, we shall ask that the rates at which blocks move and merge do not depend on the number or labels of these individuals. Hence, these rates will only depend on the collection $\{k_1, \dots, k_p, 0, \dots, 0\}$ giving the numbers of blocks contained in the different components of ξ. In order to describe the possible transitions, we need the following definition.

DEFINITION 4.9. *Let* $\hat{k} = \{k_1, \dots, k_p\}$ *and* $\hat{k}' = \{k'_1, \dots, k'_q\}$ *be two collections of (non-zero) integers. We shall write* $\hat{k} \rhd \hat{k}'$ *if* $q \geq p$, $\sum_{i=1}^q k'_i \leq \sum_{j=1}^p k_j$, *and we can arrange the elements of* $\hat{k}'$ *so that for each* $i \in \{1, \dots, p\}$, *we have* $1 \leq k'_i \leq k_i$ *and at least one of such* k'_i *is strictly less than* k_i.

Note that no collection $\hat{k}$ of integers satisfies $\{1, \dots, 1\} \rhd \hat{k}$.

For all pairs $(\hat{k}, \hat{k}')$ such that $\hat{k} \rhd \hat{k}'$, let $\vartheta_{\hat{k},\hat{k}'} \in \mathbf{R}_+$. In addition, if $\zeta \in \mathrm{P}_n^s$, let $\hat{k}(\zeta)$ be the collection of integers which gives the number of blocks within each non-empty component of ζ. Define the infinitesimal rate $q_{(\xi)}(\eta|\zeta)$ of a particular transition $\zeta \to \eta$ (when $\eta \neq \zeta$ and both belong to P_n^s) as:

- $q_{(\xi)}(\eta|\zeta) = \vartheta_{\hat{k}(\zeta),\hat{k}(\eta)}$, if η can be obtained from ζ by first merging some number (possibly zero) of blocks contained in the same component of ζ, and then moving some blocks to formerly empty demes with the restriction that only blocks originating from the same deme can be gathered into the same destination deme (again, we allow the number of blocks moved to be zero). In this case, we easily see that we must have $\hat{k}(\zeta) \rhd \hat{k}(\eta)$.
- $q_{(\xi)}(\eta|\zeta) = 0$ otherwise.

In the following, we shall assume that for any $\zeta \in \mathrm{P}_n^s$ containing more than one block in at least one component, the rates satisfy the condition

$$\sum_{\eta \in \mathrm{P}_n^s} q_{(\xi)}(\eta|\zeta) > 0.$$

These conditions ensure that, whenever a deme contains more than one lineage, a scattering or a coalescence event will happen in the future with probability one. Recall the definition of Π_n given in (4.1). From the form of the rates given above, we see that any $\eta \in \Pi_n$ is an absorbing state for ξ. Moreover, we have the following result, saying in essence that the process ξ with values in P_n^s reaches a final state in a finite number of steps, and this final state is a random variable with values in Π_n.

LEMMA 4.10. *Let* τ_π *be the stopping time defined by* $\tau_\pi := \inf\{t \geq 0 : \xi_t \in \Pi_n\}$. *Then,* τ_π *is a.s. finite and for all* $t \geq \tau_\pi$, $\xi_t = \xi_{\tau_\pi}$.

Proof. From our assumptions on the rates, the only absorbing states of the process ξ are the structured partitions contained in Π_n. Moreover, any transition results in a coarsening of the corresponding unstructured partition or in the movement of some lineages to different empty demes, so the number of transitions for ξ, starting at any $\xi_0 \in \mathrm{P}_n^s$ is bounded by n. Since the time between two events is exponentially distributed with a non-zero parameter (the sum of the rates of the possible transitions) as long as the process has not reached an absorbing state, the finiteness of the number of transitions undergone by ξ imposes that τ_π is a.s. finite. □

Let us introduce the following notation, justified by the result of Lemma 4.10.

NOTATION 4.11. *If $\zeta \in \mathrm{P}_n^s$, let $\underline{\zeta}$ denote a random variable with values in Π_n, whose distribution is that of the final state of the structured genealogical process ξ started at ζ.*

We end this subsection with the following lemma, whose main purpose is to introduce the notion of consistency for structured genealogical processes. If $\zeta \in \mathrm{P}_k^s$ and $\tilde{\zeta} \in \mathrm{P}_{k+1}^s$ for some $k \in \mathbb{N}$, let us write $\zeta \prec \tilde{\zeta}$ if the projection of $\tilde{\zeta}$ onto P_k^s (the k-tuple describing the structured partition of $1, \dots, k$) equals ζ.

LEMMA 4.12. *Suppose that ξ is defined on P_k^s for every $k \in \mathbb{N}$. The following conditions are equivalent:*

(i) For each $k \geq 1$, $\zeta, \eta \in \mathrm{P}_k^s$ and $\tilde{\zeta} \in \mathrm{P}_{k+1}^s$ such that $\zeta \prec \tilde{\zeta}$,

$$q_{(\xi)}(\eta|\zeta) = \sum_{\tilde{\eta}} q_{(\xi)}(\tilde{\eta}|\tilde{\zeta}),$$

where the sum is over all $\tilde{\eta} \in \mathrm{P}_{k+1}^s$ such that $\eta \prec \tilde{\eta}$.

(ii) The process ξ is consistent in the sense that, for all $k \geq 1$, if $\zeta \in \mathrm{P}_k^s$ and $\zeta' \in \mathrm{P}_{k+1}^s$ satisfy $\zeta \prec \zeta'$, then the law of the restriction to P_k^s of the process ξ started at ζ' is the same as the law of ξ started at ζ.

In particular, if both conditions are fulfilled and if $\eta \in \Pi_k$ has r blocks, then

$$\mathbf{P}\big[\, \underline{\zeta} = \eta \big] = \sum_{j=1}^{r+1} \mathbf{P}\big[\, \underline{\zeta'} = \eta^{(j)} \big],$$

where for each $j \in \{1, \dots, r\}$, $\eta^{(j)} \in \Pi_{k+1}$ is obtained from η by adding an empty $(k+1)$-st component to η (which turns it into an element η' of Π_{k+1}), and adding $k+1$ in the j'th block of η'. Likewise, $\eta^{(r+1)}$ is obtained by adding the singleton $\{k+1\}$ in an empty component of η'.

Proof. Let $\xi_{(k)}$ (resp. $\xi_{(k+1)}$) denote the process ξ started at $\zeta \in \mathrm{P}_k^s$ (resp. $\zeta' \in \mathrm{P}_{k+1}^s$), and call $\xi'_{(k)}$ the projection onto P_k^s of $\xi_{(k+1)}$. Since we work with finite state spaces and discrete jump processes, (ii) is equivalent to the fact that for all $\gamma, \eta \in \mathrm{P}_k^s$ the rate at which $\xi'_{(k)}$ jumps from γ to η is equal to the corresponding transition rate for $\xi_{(k)}$. By construction, the former is the sum of the rates of all the transitions from the current state of $\xi_{(k+1)}$ to a state η' such that $\eta \prec \eta'$, and so (ii) holds if and only if (i) does.

The second part of Lemma 4.12 is a direct consequence of the consistency of the process. □

2.3. Limiting process on the slow time scale. Let us now describe the form that we would expect the genealogical process to take on the slow time scale as the number of demes tends to infinity. This process $\mathcal{A}$ will have values in Π_n, so once again we construct it by specifying its transition rates.

Recall the two ingredients of the description of the structured genealogical processes indexed by $D < \infty$, given in Section 1.4. Coalescence and movement of blocks to formerly empty demes are the two kinds of events that constitute the fast process ξ, and we saw in Lemma 4.10 that the final state reached by ξ belongs to Π_n a.s. Therefore, we now need to describe how the resulting geographically separated lineages are gathered into identical demes and, potentially, merge during the same event. As in the definition of ξ, the rate at which such an event occurs will only depend on the number r of demes containing at least one lineage just after the event, on the numbers $k_1,\dots,k_r$ of blocks brought together into these components, and on the number and sizes of the groups of blocks ending up in the same demes which subsequently merge into a bigger block. Hence, we shall use the following terminology.

DEFINITION 4.13. *Let $k\geq 2$, and $k_1,\dots,k_r\geq 1$ such that $\sum_{i=1}^r k_i=k$ and at least one of the k_i's is greater than* 1. *Let also $L_1=\{l_{1,1},\dots,l_{1,i_1}\}$, …, $L_r=\{l_{r,1},\dots,l_{r,i_r}\}$ be r (unordered) sets of integers such that for $j\in\{1,\dots,r\}$, we have $\sum_{u=1}^{i_j} l_{j,u}=k_j$. We call an event in which k lineages spread in k different demes become grouped into k_1 lineages in one deme, k_2 lineages in another deme, …, and for all $j\in\{1,\dots,r\}$, $l_{j,1}$ lineages in deme j merge into one, $l_{j,2}$ into another, and so on (all mergers occur between lineages which landed in the same deme) a $(k;k_1,\dots,k_r;L_1,\dots,L_r)$*-geographical collision.

REMARK 4.14. *A geographical collision is to be understood as a particular transition. Because the order of $k_1,\dots,k_r$ does not matter, a $(k;k_1,\dots,k_r;L_1,\dots,L_r)$-geographical collision is also a $(k;k_{\sigma(1)},\dots,k_{\sigma(r)};\,L_{\sigma(1)},\dots,L_{\sigma(r)})$-geographical collision for any permutation σ of $\{1,\dots,r\}$. Thus, for a given $(k;k_1,\dots,k_r;L_1,\dots,L_r)$, the number of $(k;k_1,\dots,k_r;L_1,\dots,L_r)$-geographical collisions is*

$$\mathrm{A}(k;k_1,\dots,k_r)\prod_{m=1}^{r}\mathrm{A}(k_m;l_{m,1},\dots,l_{m,i_m}),$$

where if $k,k_1,\dots,k_r$ are such that $\sum_{i=1}^r k_i=k$ and b_j is the number of k_i's equal to j, then

$$\mathrm{A}(k;k_1,\dots,k_r)=\binom{k}{k_1,\dots,k_r}\frac{1}{\prod_{j=1}^{k}b_j!}. \tag{4.3}$$

Indeed, the binomial term gives the number of ways of choosing k_1 blocks to form a family numbered 1, *k_2 other blocks to form family n^{o}2, and so on. But any permutation of the labels of families having the same size gives the same unordered structured partition, hence the normalization by the fraction in the right-hand side of (4.3).*

Let us now define the structured genealogical process $\mathcal{A}$. The relation between the coefficients of $\mathcal{A}$ and the sequence of structured genealogical processes will be given in the next section, and we simply give the form of the limiting process here. Recall that $|\zeta|_a=k$ if the a'th component of ζ contains k blocks, and write $|\zeta|$ for the total number of blocks of $\zeta\in\mathrm{P}_n^s$, that is $|\zeta|=\sum_{a=1}^n|\zeta|_a$. Furthermore, let $\underline{\zeta}$ be a Π_n-valued random variable with the distribution specified in Notation 4.11.

DEFINITION 4.15. *For all integers and sets k, k_i and L_i satisfying the conditions of Definition 4.13, let $\lambda^g_{k;k_1,\dots,k_r;L_1,\dots,L_r}\geq 0$. Then, $(\mathcal{A}_t,t\geq 0)$ is the Markov process with values in Π_n which evolves as follows: when $\mathcal{A}_t=\chi\in\Pi_n$, any $(|\chi|;k_1,\dots,k_r;L_1,\dots,L_r)$-geographical collision occurs at rate $\lambda^g_{|\chi|;k_1,\dots,k_r;L_1,\dots,L_r}$. Given that this collision has outcome $\zeta\in\mathrm{P}_n^s$, the new value of $\mathcal{A}$ is drawn from the distribution of $\underline{\zeta}$.*

We can recover the expression for the rate of any given transition in the form

$$q(\eta|\chi)=\sum_{\zeta\in\mathrm{P}_n^s}\lambda^g_{|\chi|;k_1,\dots,k_r;L_1,\dots,L_r}\underline{\zeta}[\eta],$$

where the rate $\lambda^g_{|\chi|;k_1,\dots,k_r;L_1,\dots,L_r}$ in the term of the sum labeled by a given ζ is the rate of occurrence of the only possible geographical collision turning χ into ζ, if such a collision exists. If it does not, we set the rate to 0. Consequently, the previous description does specify a Markov process on Π_n.

Observe that this description allows ‘ghost events’ in which lineages are gathered in identical demes by a geographical collision and then scattered again in different demes without coalescing, so that the actual transition is of the form $\chi \to \chi$. However, we shall need to keep track of these ghost events in the proof of convergence of the structured genealogical processes. Therefore, we shall always consider them as events which do occur at a certain rate but have no effect on $\mathcal{A}$. To finish the description of our limiting process, we have the following result, which in fact describes the unstructured genealogy under some additional conditions.

PROPOSITION 4.16. *For each $\zeta \in \Pi_n$, let us define $\zeta^u \in \mathcal{P}_n$ as the unstructured partition of n induced by ζ. Then the unstructured genealogical process $(\mathcal{A}^u_t, t \geq 0)$ induced by $\mathcal{A}$ is a Markov process with values in $\mathcal{P}_n$. Suppose in addition that Condition (i) of Lemma 4.12 holds and that the λ^g's satisfy the following consistency equations: for all $k \in \mathbb{N}$ and compatible $k_1, \ldots, k_r, L_1, \ldots, L_r$,*

(4.4)

$$\lambda^g_{k;k_1,\ldots,k_r;L_1,\ldots,L_r} = \sum_{u=1}^{r}\sum_{j=1}^{i_u+1} \lambda^g_{k+1;k_1,\ldots,k_u+1,\ldots,k_r;L_1,\ldots,L_u^{(j)},\ldots,L_r} + \lambda^g_{k+1;k_1,\ldots,k_r,1;L_1,\ldots,L_r,\{1\}},$$

where for each $u \leq r$

$$L_u^{(j)} = \begin{cases} \{l_{u,1}, \ldots, (l_{u,j}+1), \ldots, l_{u,i_u}\} & \text{if } j \leq i_u \\ \{l_{u,1}, \ldots, l_{u,i_u}, 1\} & \text{if } j = i_u + 1. \end{cases}$$

(In particular, if instantaneous coalescence after the gathering of lineages is forbidden, then the λ^g's are associated to a Ξ-coalescent.)

Then $(\mathcal{A}^u_t, t \geq 0)$ is the restriction to $\mathcal{P}_n$ of a Ξ-coalescent on the partitions of $\mathbb{N}$.

REMARK 4.17. *By fixing $k, k_1, \ldots, k_r$ and summing over all compatible integer sets $L_1, \ldots, L_r$, we see that the rates $\tilde{\lambda}^g_{k;k_1,\ldots,k_r}$ at which k lineages lying in k different demes end up in a configuration where k_1 lineages are in the same deme, k_2 in another one, and so on (regardless of how many of them merge instantaneously thereafter), are associated to a Ξ-coalescent whenever Condition (4.4) holds.*

Proof. Any component of $\gamma \in \Pi_n$ contains at most one block and all n-tuples are defined up to a permutation of their components, so the map $\Pi_n \to \mathcal{P}_n : \gamma \mapsto \gamma^u$ is a measurable bijection between Π_n and $\mathcal{P}_n$. Thus, $\mathcal{A}^u$ inherits the Markov property of $\mathcal{A}$, and its transition rates $q^u(\eta^u|\,\gamma^u)$ are obviously given by $q^u(\eta^u|\,\gamma^u) = q(\eta|\,\gamma)$.

Let us turn to the second part of Proposition 4.16. By assumption, $\mathcal{A}^u$ only coarsens as time goes on and it is easy to check that all transition rates $\rho(k; k_1, \ldots, k_r)$ from a partition with k blocks to a partition obtained by merging k_1 of those blocks into one, k_2 into a second one, ... ($k_1, \ldots, k_r \in \mathbb{N}$, $\sum_{i=1}^r k_i = k$), are equal and depend only on $k, k_1, \ldots, k_r$ (the order of $k_1, \ldots, k_r$ does not matter). Therefore, we need only check the consistency condition given in [**Sch00**] to identify $\mathcal{A}^u$ as the restriction to the partitions of $[n]$ of a Ξ-coalescent. As the rates do not depend on n, let us rather work in Π_k with $\gamma = (\{\{1\}\}, \ldots, \{\{k\}\})$ and $\eta = (\{\{1, \ldots, k_1\}\}, \{\{k_1+1, \ldots, k_1+k_2\}\}, \ldots, \{\{k_1+\ldots+k_{r-1}+1, \ldots, k\}\}, \emptyset, \ldots, \emptyset)$, and check that

$$\rho(k; k_1, \ldots, k_r) = \sum_{i=1}^{r} \rho(k+1; k_1, \ldots, k_i+1, \ldots, k_r) + \rho(k+1; k_1, \ldots, k_r, 1).$$

Since the λ^g's satisfy (4.4), we have

$\rho(k\,; k_1, \ldots, k_r) = q(\eta|\,\gamma)$

$$= \sum_{\zeta\in \mathrm{P}_k^s} \lambda^g_{k;l_1,\dots,l_s;L_1,\dots,L_s}\, \underline{\zeta}[\eta]$$

$$= \sum_{\zeta\in \mathrm{P}_k^s}\sum_{v=1}^{s}\sum_{j=1}^{i_v+1} \lambda^g_{k+1;l_1,\dots,l_v+1,\dots,l_s;L_1,\dots,L_v^{(j)},\dots,L_s}\, \underline{\zeta}[\eta] + \sum_{\zeta\in \mathrm{P}_k^s} \lambda^g_{k+1;l_1,\dots,l_s,1;L_1,\dots,L_s,\{1\}}\underline{\zeta}[\eta]. \tag{4.5}$$

We wish to compare this rate to the rates corresponding to $k+1$ blocks. To this end, let us define $\zeta_{(v,j)} \in \mathrm{P}_{k+1}^s$ for all $\zeta \in \mathrm{P}_k^s$ with l non-empty components and $v \in \{1,\dots,l+1\}$, $j \in \{1,\dots,i_v+1\}$ (i_v being the number of blocks in the v'th non-empty component of ζ) by turning ζ into a $(k+1)$-tuple and adding individual $k+1$ in the j'th block of the v'th component of the new vector ($v = l+1$ means that we add the block $\{k+1\}$ in the extra component, and likewise $j = i_v+1$ means that we add the block $\{k+1\}$ in the v'th component of the new vector). For example, with the previous notation γ,

$$\gamma_{(1,2)} = (\{\{1\},\{k+1\}\},\dots,\{\{k\}\},\emptyset) \quad \text{and} \quad \gamma_{(k+1,1)} = (\{\{1\}\},\dots,\{\{k\}\},\{\{k+1\}\}).$$

Define also $\gamma^{(j)} \in \Pi_{k+1}$, for all $\gamma \in \Pi_k$ with r blocks and $j \in \{1,\dots,r+1\}$, by turning γ into a $(k+1)$-tuple and adding individual $k+1$ in the block of the j'th component of the new vector. Once again, $j = k+1$ means that we add a block $\{k+1\}$ in the extra component. For instance,

$$\eta^{(1)} = (\{\{1,\dots,k_1,k{+}1\}\},\{\{k_1{+}1,\dots,k_1{+}k_2\}\},\dots,\{\{k_1{+}\dots{+}k_{r-1}{+}1,\dots,k\}\},\emptyset,\dots,\emptyset)$$

and

$$\begin{aligned}\eta^{(k+1)} = (\{\{1,\dots,k_1\}\},\{\{k_1+1,\dots,k_1+k_2\}\},\dots,\{\{k_1+\dots+k_{r-1}+1,\dots,k\}\},&\\ \{\{k+1\}\},\emptyset,\dots,\emptyset).&\end{aligned}$$

With this notation, we see that

$$\sum_{i=1}^{r} \rho(k+1;k_1,\dots,k_i+1,\dots,k_r) + \rho(k+1;k_1,\dots,k_r,1) = \sum_{i=1}^{r+1} q(\eta^{(i)} | \,\gamma_{(k+1,1)}). \tag{4.6}$$

For all $\zeta' \in \mathrm{P}_{k+1}^s$, there exists a unique triplet (ζ,v,j) where $\zeta \in \mathrm{P}_k^s$ has l non-empty components, $v \in \{1,\dots,l+1\}$ and $j \in \{1,\dots,i_v+1\}$ such that $\zeta' = \zeta_{(v,j)}$. Indeed, ζ is given by the partition of $\{1,\dots,k\}$ induced by ζ', v is the component containing $k+1$ and j is the block of that component in which $k+1$ lies. Therefore, the right-hand side of (4.6) is equal to

$$\begin{aligned}&\sum_{i=1}^{r+1}\sum_{\zeta\in \mathrm{P}_k^s}\sum_{v=1}^{s}\sum_{j=1}^{i_v+1} \lambda^g_{k+1;l_1,\dots,l_v+1,\dots,l_s;L_1,\dots,L_v^{(j)},\dots,L_s}\, \underline{\zeta_{(v,j)}}[\eta^{(i)}]\\ &\qquad + \sum_{i=1}^{r+1}\sum_{\zeta\in \mathrm{P}_k^s} \lambda^g_{k+1;l_1,\dots,l_s,1;L_1,\dots,L_s,\{1\}}\, \underline{\zeta_{(s+1,1)}}[\eta^{(i)}],\end{aligned} \tag{4.7}$$

where s and the coefficients $\lambda^g_{k+1;l_1,\dots,l_v+1,\dots,l_s;L_1,\dots,L_v^{(j)},\dots,L_s}$ correspond to the particular ζ indexing the term of the sum. Let us look at a particular ζ in the second sum. The block $\{k+1\}$ remains a singleton just after the geographical collision, so it is not affected by a following genealogical event and $\underline{\zeta_{(s+1,1)}}[\eta^{(i)}] = 0$ for all $i \in \{1,\dots,r\}$. Lemma 4.12 hence implies that $\underline{\zeta_{(s+1,1)}}[\eta^{(r+1)}] = \underline{\zeta}[\eta]$, and the second term of (4.7) is equal to

$$\sum_{\zeta\in \mathrm{P}_k^s} \lambda^g_{k+1;l_1,\dots,l_s,1;L_1,\dots,L_s,\{1\}}\, \underline{\zeta}[\eta].$$

Let us look at a particular ζ in the first sum, now. When $v \leq s$, the corresponding geographical collision brings $k+1$ in a block of the v'th component of ζ. By the second part of Lemma 4.12, the probability that the final state of all the blocks different from $k+1$ is given by η is equal to the sum over all corresponding final states of these blocks and $k+1$. But taking the sum over i in $\sum_{i=1}^{s+1} \underline{\zeta_{(v,j)}}[\eta^{(i)}]$ boils down to considering all such final states, since $\underline{\zeta_{(v,j)}}[\eta^{(i)}] = 0$ if the individuals in the i'th block of η were not in the v'th component of ζ before their rearrangement by the genealogical process (recall that, under the action of ξ, lineages can merge only if they start in the same deme). Therefore, we obtain that, for all $\zeta \in \mathrm{P}_k^s$ and compatible v, j,

$$\sum_{i=1}^{r+1} \underline{\zeta_{(v,j)}}[\eta^{(i)}] = \underline{\zeta}[\eta].$$

Coming back to expressions (4.6) and (4.7), we obtain

$$\begin{aligned}
&\sum_{i=1}^{r} \rho\;(k+1; k_1, \dots, k_i+1, \dots, k_r) + \rho(k+1; k_1, \dots, k_r, 1) \\
&\quad = \sum_{\zeta \in \mathrm{P}_k^s} \sum_{v=1}^{s} \sum_{j=1}^{i_v+1} \lambda^g_{k+1; l_1, \dots, l_v+1, \dots, l_s; L_1, \dots, L_v^{(j)}, \dots, L_s} \, \underline{\zeta}[\eta] + \sum_{\zeta \in \mathrm{P}_k^s} \lambda^g_{k+1; l_1, \dots, l_s, 1; L_1, \dots, L_s, \{1\}} \, \underline{\zeta}[\eta] \\
&\quad = \rho(k; k_1, \dots, k_r),
\end{aligned}$$

where the last equality follows from (4.5). This completes the proof of Proposition 4.16. □

3. Convergence of the structured genealogical processes

Now that we have constructed the potential limits for our sequence of structured genealogical processes on the fast and slow time scales, let us state precisely what conditions we impose and in which sense $\mathcal{A}^D_{r_D^{-1}\cdot}$ and $\mathcal{A}^D$ converge.

3.1. Description of the conditions. Let $n \geq 1$ be the sample size and define two types of events:

- Type 1: some lineages contained in the same demes merge and some move (potentially in groups) to empty islands. The number of lineages involved in either step can be zero (meaning that only coalescence or only scattering has occurred), and lineages starting from different demes are not gathered into the same deme by the event.
- Type 2: k lineages move, but at least one of them lands in a non-empty deme or at least two dispersing lineages not coming from the same deme are gathered. During that event, k_1 lineages end up in the same deme, k_2 lineages in another, and so on. This is immediately followed by the coalescence of some lineages lying in identical demes (the number of such mergers can be zero, meaning that the lineages have only moved).

By our assumptions on the genealogical processes, these two types describe all kinds of events which can happen to the structured genealogical process $\mathcal{A}^D$, for each D. For conciseness, we shall call an event of type i an i-event. Assume now that, when $\mathcal{A}^D$ has value $\zeta \in \mathrm{P}_n^s$ and $\eta \in \mathrm{P}_n^s$ is a possible new value compatible with the type of the event (o's hold as D goes to infinity):

1. The rate of occurrence of a particular 1-event $\zeta \to \eta$ can be written

$$r_D \vartheta_{\hat{k}(\zeta), \hat{k}(\eta)} + \nu^{(n)}(\zeta, \eta) + o(1) \qquad \text{as } D \to \infty,$$

where for each n, $\nu^{(n)}(\cdot, \cdot)$ is a bounded function on $(\mathrm{P}_n^s)^2$ and $r_D \to \infty$ as $D \to \infty$.

2. Consider a 2-event involving k lineages, for which there exist $k_1, \dots, k_r \geq 1$ such that $\sum_{i=1}^{r} k_i = |\zeta|$ and there exist r sets of integers $L_1 = \{l_{1,1}, \dots, l_{1,i_1}\}$, $\dots$, $L_r = \{l_{r,1}, \dots, l_{r,i_r}\}$ such that for all $j \in \{1, \dots, r\}$ we have $\sum_{u=1}^{i_j} l_{j,u} = k_j$, satisfying: in the new structured partition, k_1 lineages end up in one deme, k_2 in another deme, ... , and for all $j \in \{1, \dots, r\}$, $l_{j,1}$ lineages in deme j merge into one, l_{j,i_2} into another one, and so on (once again, all mergers occur between lineages lying in the same deme). Then the rate of occurrence of any such event is of the form
$$l_k^{(n)}(\zeta, \eta) + o(1),$$
where for each n and all $k \leq n$, $l_k^{(n)}$ is a bounded function on $\mathrm{P}_n^s \times \mathrm{P}_n^s$, and in particular if $\zeta \in \Pi_n$,
$$l_k^{(n)}(\zeta, \eta) = \lambda^g_{|\zeta|; k_1, \dots, k_r; L_1, \dots, L_r}.$$
Here again, the order of $k_1, \dots, k_r$ does not matter.
3. The ϑ's correspond to a structured genealogical process ξ as described in the last section, and the λ^g's satisfy the consistency equations (4.4).

Morally, the coalescence of lineages occupying common demes and the scattering of such lineages into empty demes occur more and more rapidly as D tends to infinity, whereas events collecting lineages into common demes occur at bounded rates. Other events are less and less frequent, so that in the limit we obtain a separation of time scales between the instantaneous structured genealogical process and the slow collecting phase of the limiting unstructured genealogical process. Notice that 1-events do not affect a structured partition contained in Π_n.

Let $G^{n,D}$ denote the generator of the genealogical process of a sample of n individuals when the number of demes is D. For each D, the domain $\mathcal{D}(G^{n,D})$ of $G^{n,D}$ contains the measurable symmetric functions of n variables (by symmetric, we mean invariant under all permutations of the variables). From the last remark, we see that for all $f \in \mathcal{D}(G^{n,D})$, the parts of $G^{n,D}f$ corresponding to 1-events vanish on Π_n. Furthermore, we can define linear operators Ψ^n, Γ^n and R^n_D such that $G^{n,D}$ has the following form:
$$G^{n,D} = r_D \Psi^n + \Gamma^n + R^n_D.$$
More precisely, for every function f as above and each $\zeta \in \mathrm{P}_n^s$, we have
$$\Psi^n f(\zeta) = \sum_{\eta \in \mathrm{P}_n^s} \vartheta_{\hat{k}(\zeta), \hat{k}(\eta)} \big(f(\eta) - f(\zeta)\big),$$
$$\Gamma^n f(\zeta) = \sum_{\eta \in \mathrm{P}_n^s} \big\{\nu^{(n)}(\zeta, \eta) + l_k^{(n)}(\zeta, \eta)\big\} \big(f(\eta) - f(\zeta)\big),$$
and by the nonnegativity of their coefficients, these two operators can each be viewed as generating a jump process independent of D. In particular, we can define the structured genealogical process ξ on P_n^s as the process generated by Ψ^n. The remaining terms $o(1)$ in Assumptions 1 and 2 constitute the coefficients of the (not necessarily positive) operator R^n_D, and so if we again use the operator norm introduced in (4.2), the finiteness of the number of possible transitions guarantees that $\langle R^n_D \rangle = o(1)$ as $D \to \infty$.

3.2. Convergence of the structured genealogical processes. The main result of this section is the convergence of the finite-dimensional distributions of the P_n^s-valued structured genealogical processes $\mathcal{A}^D$ to the corresponding ones of $\mathcal{A}$, except at time $t = 0$. The difficulty stems from the fact that the sequence of generators $G^{n,D}$ is unbounded because of the fast genealogical events driven by Ψ^n. The proof consists in essence in showing that the dynamics of the genealogical processes become very close to the description of the dynamics of $\mathcal{A}$, in that for D large enough, once a Γ^n-event (i.e., a geographical collision) occurs, enough Ψ^n-events happen in a very short period of time to bring the structured partition back into Π_n. During that short period, the probability that a Γ^n- or an R^n_D-event occurs is vanishingly small so that at the time when $\mathcal{A}^D$ re-enters Π_n, with a high probability it has the distribution of the final state of ξ started at the structured partition created by the geographical collision. Overall, R^n_D-events are more and more infrequent and do not occur in the limit.

Before stating the results of this section, let us define the probability measures of interest. We take for granted the fact that the processes ξ and $\mathcal{A}^D$ for each $D \in \mathbb{N}$ and all $n \in \mathbb{N}$ can be constructed on the same probability space $(\Omega, \mathbf{P}, \mathcal{F})$. For all $\zeta \in \mathrm{P}^s_n$, we thus denote the probability measure under which these processes start at ζ by $\mathbf{P}_\zeta$. Likewise, let $(\Omega', \mathbb{P}, \mathcal{F}')$ be the probability space on which the processes $\mathcal{A}$ and χ (see Definition 4.19) are defined for all $n \in \mathbb{N}$. $\mathbb{P}_\eta$ denotes the probability measure under which these processes start at $\eta \in \Pi_n$.

With this notation, Theorem 4.4 can be restated as:

Theorem 4.4’. *Suppose that the conditions stated in Section 3.1 hold, and let $\zeta \in \mathrm{P}^s_n$. Then, the structured genealogical processes $\mathcal{A}^D$ started at ζ converge to the process $\mathcal{A}$ started at $\underline{\zeta}$ as D tends to infinity, in the sense that for all $0 < t_1 < \ldots < t_p$,*

$$\mathbf{P}_\zeta(\mathcal{A}^D_{t_1}, \ldots, \mathcal{A}^D_{t_p}) \Rightarrow \mathbb{P}_{\underline{\zeta}}(\mathcal{A}_{t_1}, \ldots, \mathcal{A}_{t_p}) \qquad \text{as } D \to \infty,$$

where $\mathbf{P}_\zeta(X)$ stands for the law of the random variable X under $\mathbf{P}_\zeta$ and $\mathbb{P}_{\underline{\zeta}}(X)$ is defined similarly.

We also have the following result.

PROPOSITION 4.18. *Assume again that the conditions of Section 3.1 hold. Then the sequence of $D_{\mathrm{P}^s_n}[0, \infty)$-valued processes $\{\mathcal{A}^D_{r_D^{-1}t}, t \geq 0\}$ converges in distribution to the structured genealogical process ξ introduced in Section 2.2.*

The proof of Proposition 4.18 is a direct consequence of the uniform convergence of the generator of $\mathcal{A}^D_{r_D^{-1}\cdot}$ (namely $r_D^{-1}G^{n,D}$) to the generator of ξ and the finiteness of the state space. A coupling with ξ shows that the first time at which both processes differ when started from the same value tends to infinity in probability, which is the main argument to obtain the desired convergence. The proof being immediate, we turn instead to the proof of Theorem 4.4.

Let us first introduce the following notation, for each $D \in \mathbb{N}$:

$$\sigma^D_1 := \inf\{t \geq 0 : \mathcal{A}^D_t \in \Pi_n\}, \qquad \tau^D_1 := \inf\{t > \sigma^D_1 : \text{ a 2−event occurs at } t\},$$

and for all $i \geq 2$,

$$\sigma^D_i := \inf\{t \geq \tau^D_{i-1} : \mathcal{A}^D_t \in \Pi_n\}, \qquad \tau^D_i := \inf\{t > \sigma^D_i : \text{ a 2−event occurs at } t\}, \tag{4.8}$$

with the convention that $\inf \emptyset = +\infty$ and if σ^D_i or $\tau^D_i = +\infty$, then the following random times are all equal to $+\infty$. Note that if a 2-event occurs, its outcome may still be in Π_n (if all lineages gathered in identical demes merge into one lineage in each of these demes). In that case, $\sigma^D_{i+1} = \tau^D_i$. Let us also denote the ranked epochs of events occurring to the process $\mathcal{A}$ by $\sigma_i,\ i \geq 1$, including what we previously called the ‘ghost events’, with the conventions that $\sigma_1 = 0$ and $\sigma_k = +\infty$ for $k \geq j + 1$ if there are no more events after the j’th transition.

Proof of Theorem 4.4. We start by proving the convergence of the one-dimensional distributions, then establish the convergence of the finite dimensional distributions by induction on their dimension. Since the sample size is fixed, we drop the superscript n in our notation.

As a first step, let us state the following definition and two lemmas, which will be useful in the course of the proof. For the sake of clarity, the proofs of the lemmas are postponed until after the proof of Theorem 4.4.

DEFINITION 4.19. *Let $(\chi_t, t \in [0,T))$ denote a Π_n-valued Markov process generated by Γ^n, where T is defined as*

$$T := \inf\{t \geq 0 : \ \chi_t \notin \Pi_n\}.$$

Then, for all $\eta \in \Pi_n$, $\chi(\eta)$ is defined as a P_n^s-random variable distributed like the outcome of the first geographical collision when χ starts at η (this event is always defined if the λ^g's satisfy (4.4) and η has at least two blocks, since the coefficients $\tilde{\lambda}^g$ are the rates of a Ξ-coalescent as mentioned in Remark 4.17).

LEMMA 4.20. *Let $i \geq 1$. Then for all bounded measurable functions f on $\mathbf{R}_+ \times \mathrm{P}_n^s$, we have*

$$\lim_{D\to\infty} \mathbf{E}_\zeta\big[f(\sigma_i^D, \mathcal{A}^D_{\sigma_i^D})\,\mathbb{I}_{\{\sigma_i^D<\infty\}}\big] = \mathbb{E}_{\underline{\zeta}}\big[f(\sigma_i, \mathcal{A}_{\sigma_i})\,\mathbb{I}_{\{\sigma_i<\infty\}}\big],$$
$$\lim_{D\to\infty} \mathbf{E}_\zeta\big[f(\tau_i^D, \mathcal{A}^D_{\tau_i^D})\,\mathbb{I}_{\{\tau_i^D<\infty\}}\big] = \mathbb{E}_{\underline{\zeta}}\big[f\big(\sigma_{i+1}, \chi(\mathcal{A}_{\sigma_i})\big)\,\mathbb{I}_{\{\sigma_{i+1}<\infty\}}\big].$$

In particular, by taking $f(t,\eta) = \mathbb{I}_{\{t\leq s\}}$ for all $s > 0$, we obtain that the law under $\mathbf{P}_\zeta$ of the $[0,+\infty]$-valued random variable σ_i^D (resp. τ_i^D) converges to the law under $\mathbb{P}_{\underline{\zeta}}$ of σ_i (resp. σ_{i+1}).

LEMMA 4.21. *Let $t \in (0,\infty)$ and let $i \in \mathbb{N}$ be such that $\mathbb{P}_{\underline{\zeta}}[\sigma_i < \infty] > 0$. By Lemma 4.20, we also have for D large enough $\mathbf{P}_\zeta[\sigma_i^D < \infty] > 0$. Let f be a real-valued function on P_n^s. Then*

$$\lim_{D\to\infty} \mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\,\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t) \bigm| \sigma_i^D < \infty\big] = \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\,\mathbb{I}_{[\sigma_i,\sigma_{i+1})}(t) \bigm| \sigma_i < \infty\big].$$

Fix $t > 0$, let f be a real-valued function on P_n^s and denote the supremum norm of f by $\|f\|$. We have for each D and all $N \in \mathbb{N}$:

$$\begin{aligned}
&\Big|\mathbf{E}_\zeta\,[f(\mathcal{A}_t^D)] - \mathbb{E}_{\underline{\zeta}}[f(\mathcal{A}_t)]\Big| \\
&= \Bigg|\mathbf{E}_\zeta\Bigg[\sum_{i=1}^{\infty} f(\mathcal{A}_t^D)\,\mathbb{I}_{[\tau_{i-1}^D,\sigma_i^D)}(t) + \sum_{i=1}^{\infty} f(\mathcal{A}_t^D)\,\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\Bigg] - \mathbb{E}_{\underline{\zeta}}\Bigg[\sum_{i=1}^{\infty} f(\mathcal{A}_t)\,\mathbb{I}_{[\sigma_i,\sigma_{i+1})}(t)\Bigg]\Bigg| \\
&\leq \sum_{i=1}^{N}\Big|\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\,\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\big] - \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\,\mathbb{I}_{[\sigma_i,\sigma_{i+1})}(t)\big]\Big| \\
&\quad + \sum_{i=1}^{N}\mathbf{E}_\zeta\big[|f(\mathcal{A}_t^D)|\,\mathbb{I}_{[\tau_{i-1}^D,\sigma_i^D)}(t)\big] + \Big|\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\mathbb{I}_{\{t\geq\tau_N^D\}}\big]\Big| + \Big|\mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\mathbb{I}_{\{t\geq\sigma_{N+1}\}}\big]\Big|,
\end{aligned} \tag{4.9}$$

where $\tau_0 := 0$. Let $\epsilon > 0$. The random variables σ_i are the jump times of $\mathcal{A}$, the rates of which are bounded above by a constant $b \geq 0$. Thus, for each $N \geq 1$, σ_N is bounded below by the sum of N independent exponentials with parameter b, and so there exists $N \geq 1$ such that

$$\mathbb{P}_{\underline{\zeta}}[\sigma_{N+1} < t] < \frac{\epsilon}{4\|f\|}\,.$$

In addition, $\tau_N^D \Rightarrow \sigma_{N+1}$ by Lemma 4.20, so there exists a D_0 such that for $D \geq D_0$,

$$\mathbf{P}_\zeta[\tau_{N+1}^D < \infty] < \frac{\epsilon}{4\|f\|}.$$

Consequently, for $D \geq D_0$ we have

$$\Big|\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\mathbb{I}_{\{t\geq\tau_N^D\}}\big]\Big| + \Big|\mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\mathbb{I}_{\{t\geq\sigma_{N+1}\}}\big]\Big| \leq \|f\| \Big(\mathbf{P}_\zeta[\tau_N^D \leq t] + \mathbb{P}_{\underline{\zeta}}[\sigma_{N+1}\leq t]\Big) < \frac{\epsilon}{2}. \tag{4.10}$$

Let $i \in \{1,\dots,N\}$. We have

$$\mathbf{E}_\zeta\big[\big|f(\mathcal{A}_t^D)\big|\,\mathbb{I}_{[\tau_{i-1}^D,\sigma_i^D)}(t)\big] \leq \|f\|\,\mathbf{P}_\zeta\big[\tau_{i-1}^D \leq t < \sigma_i^D\big] = \|f\|\,\left(\mathbf{P}_\zeta\big[\tau_{i-1}^D \leq t\big] - \mathbf{P}_\zeta\big[\sigma_i^D \leq t\big]\right).$$

By Lemma 4.20, both τ_{i-1}^D and σ_i^D converge in law towards σ_i (whose distribution function is continuous on $\mathbf{R}_+$), so the right-hand side of the last inequality tends to 0 when $D \to \infty$. Hence, there exists a D_1 such that for all $D \geq D_1$,

$$\sum_{i=1}^{N} \mathbf{E}_\zeta\big[\big|f(\mathcal{A}_t^D)\big|\,\mathbb{I}_{[\tau_{i-1}^D,\sigma_i^D)}(t)\big] \leq \frac{\epsilon}{4}. \tag{4.11}$$

Once again, let $i \in \{1,\dots,N\}$. If $\mathbb{P}_{\underline{\zeta}}\big[\sigma_i \leq t\big] = 0$, then $\mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\,\mathbb{I}_{[\sigma_i,\sigma_{i+1})}(t)\big] = 0$ and

$$\Big|\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\,\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\big]\Big| \leq \|f\|\,\mathbf{P}_\zeta\big[\sigma_i^D \leq t\big] \to 0$$

as D tends to infinity, by Lemma 4.20 and the continuity of the distribution function of σ_i in t. If $\mathbb{P}_{\underline{\zeta}}\big[\sigma_i \leq t\big] > 0$, we also have $\mathbf{P}_\zeta\big[\sigma_i^D \leq t\big] > 0$ for D large enough, so we can write

$$\begin{aligned}
\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\,\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\big] &= \mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\,\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\big|\,\sigma_i^D < \infty\big]\mathbf{P}_\zeta[\sigma_i^D < \infty] \\
&\to \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\,\mathbb{I}_{[\sigma_i,\sigma_{i+1})}(t)\big|\,\sigma_i < \infty\big]\mathbb{P}_{\underline{\zeta}}[\sigma_i < \infty] \\
&= \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\,\mathbb{I}_{[\sigma_i,\sigma_{i+1})}(t)\big],
\end{aligned}$$

where the convergence on the second line stems from Lemma 4.21 and the convergence in distribution of σ_i^D towards σ_i. Consequently, there exists D_2 such that for all $D \geq D_2$,

$$\sum_{i=1}^{N} \Big|\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\,\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\big] - \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\,\mathbb{I}_{[\sigma_i,\sigma_{i+1})}(t)\big]\Big| \leq \frac{\epsilon}{4}. \tag{4.12}$$

Combining to (4.9), (4.10), (4.11) and (4.12), we obtain for all $D \geq \max\{D_0, D_1, D_2\}$

$$\Big|\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\big] - \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\big]\Big| \leq \epsilon.$$

We can hence conclude that

$$\lim_{D\to\infty} \mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\big] = \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_t)\big],$$

which completes the proof of the convergence of the one-dimensional distributions of $\mathcal{A}^D$ under $\mathbf{P}_\zeta$ to the corresponding ones of $\mathcal{A}$ under $\mathbb{P}_{\underline{\zeta}}$.

Let us now turn to the convergence of the finite-dimensional distributions. We prove by induction on p that, for all $0 < t_1 < \ldots < t_p$, $\mathbf{P}_\zeta(\mathcal{A}_{t_1}^D, \dots, \mathcal{A}_{t_p}^D) \Rightarrow \mathbb{P}_{\underline{\zeta}}(\mathcal{A}_{t_1}, \dots, \mathcal{A}_{t_p})$ as $D \to \infty$. By the preceding step, the case $p = 1$ is already established. Let $p \geq 2$, and suppose that the convergence holds for the $(p-1)$-dimensional distributions. Let $0 < t_1 < \ldots < t_p$, and let $f_1, \dots, f_p$ be real-valued functions on P_n^s. We denote the σ-field generated by $\big\{\mathcal{A}_s^D,\ s \in [0,t]\big\}$ by $\mathcal{F}_t^D$. Then,

$$\begin{aligned}
\mathbf{E}_\zeta\Bigg[\prod_{i=1}^{p} f_i(\mathcal{A}_{t_i}^D)\Bigg] &= \mathbf{E}_\zeta\Bigg[\mathbf{E}\Bigg[\prod_{i=1}^{p} f_i(\mathcal{A}_{t_i}^D)\Bigg|\mathcal{F}_{t_{p-1}}^D\Bigg]\Bigg] \\
&= \mathbf{E}_\zeta\Bigg[\prod_{i=1}^{p-1} f_i(\mathcal{A}_{t_i}^D)\mathbf{E}_{\mathcal{A}_{t_{p-1}}^D}\big[f_p(\tilde{\mathcal{A}}_{t_p - t_{p-1}}^D)\big]\Bigg] \qquad \text{by the Markov property}
\end{aligned}$$

$$= \mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}^D_{t_i})\Big(\sum_{\eta\in \mathrm{P}^s_n} p^D(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})f_p(\eta)\Big)\Big],$$

where here and in the following $\tilde{X}$ denotes an independent version of the random variable X, the second expectation is taken with regards to $\tilde{X}$, and $p^D(\cdot,\cdot,s)$ is the transition kernel of $\mathcal{A}^D$ corresponding to time s. Continuing the preceding equalities, we obtain

$$\begin{aligned}\mathbf{E}_\zeta\Big[\prod_{i=1}^{p} f_i(\mathcal{A}^D_{t_i})\Big] &= \sum_{\eta\in \mathrm{P}^s_n} f_p(\eta)\,\mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}^D_{t_i})\,p^D(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})\Big]\\ &= \sum_{\eta\in \mathrm{P}^s_n} f_p(\eta)\,\mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}^D_{t_i})\,p^D(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})\,\mathbb{I}_{\{\mathcal{A}^D_{t_{p-1}}\notin\,\Pi_n\}}\Big]\\ &\quad+\sum_{\eta\in \mathrm{P}^s_n} f_p(\eta)\,\mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}^D_{t_i})\,p^D(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})\,\mathbb{I}_{\{\mathcal{A}^D_{t_{p-1}}\in\,\Pi_n\}}\Big].\end{aligned}\tag{4.13}$$

For all $\eta\in\mathrm{P}^s_n$,

$$\begin{aligned}\Big|\mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}^D_{t_i})\,p^D(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})\,\mathbb{I}_{\{\mathcal{A}^D_{t_{p-1}}\notin\,\Pi_n\}}\Big]\Big| &\le \Big(\prod_{i=1}^{p-1}\|f_i\|\Big)\,\mathbf{P}_\zeta\big[\mathcal{A}^D_{t_{p-1}}\notin\Pi_n\big]\\ &\to 0\end{aligned}\tag{4.14}$$

by the convergence of $\mathcal{A}^D_{t_{p-1}}$ to $\mathcal{A}_{t_{p-1}}$ in distribution and the finiteness of P^s_n. As the sum in the first term in the right-hand side of (4.13) is finite, (4.14) implies that this sum tends to 0 when D grows to infinity. Moreover, the convergence in law of $\mathcal{A}^D_{t_p-t_{p-1}}$ to $\mathcal{A}_{t_p-t_{p-1}}$, the finiteness of P^s_n and the fact that $\underline{\gamma}=\gamma$ a.s. if $\gamma\in\Pi_n$ enable us to write

$$\max_{\gamma\in\Pi_n}\ \max_{\eta\in\mathrm{P}^s_n}\ \big|p^D(\gamma,\eta,t_p-t_{p-1})-p(\gamma,\eta,t_p-t_{p-1})\big|\to 0\qquad\text{as } D\to\infty,\tag{4.15}$$

where $p(\gamma,\eta,t_p-t_{p-1})$ is the transition kernel of $\mathcal{A}$ corresponding to time t_p-t_{p-1}, extended to $\eta\notin\Pi_n$ by $p(\gamma,\eta,t_p-t_{p-1})=0$. Now, we have for all $\eta\in\mathrm{P}^s_n$

$$\begin{aligned}&\mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i\,(\mathcal{A}^D_{t_i})\,p^D(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})\,\mathbb{I}_{\{\mathcal{A}^D_{t_{p-1}}\in\,\Pi_n\}}\Big]\\ &\quad=\mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}^D_{t_i})\,\big(p^D(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})-p(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})\big)\,\mathbb{I}_{\{\mathcal{A}^D_{t_{p-1}}\in\,\Pi_n\}}\Big]\\ &\quad\quad+\mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}^D_{t_i})\,p(\mathcal{A}^D_{t_{p-1}},\eta,t_p-t_{p-1})\,\mathbb{I}_{\{\mathcal{A}^D_{t_{p-1}}\in\,\Pi_n\}}\Big]\end{aligned}\tag{4.16}$$

The first expression in the right-hand side of (4.16) tends to 0 by (4.15) and dominated convergence. As for the second quantity in (4.16), for each $\eta\in\mathrm{P}^s_n$ the function $\gamma\mapsto p(\gamma,\eta,t_p-t_{p-1})\,\mathbb{I}_{\{\gamma\in\,\Pi_n\}}$ (vanishing on $\mathrm{P}^s_n\setminus\Pi_n$) is necessarily continuous and bounded on the finite set P^s_n, so by the induction hypothesis for $p-1$, we have

$$\begin{aligned}\lim_{D\to\infty}\mathbf{E}_\zeta\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}^D_{t_i})\,p(\mathcal{A}^D_{t_{p-1}},\eta\,,t_p-t_{p-1})\,\mathbb{I}_{\{\mathcal{A}^D_{t_{p-1}}\in\,\Pi_n\}}\Big]&\\ =\mathbb{E}_{\underline{\zeta}}\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}_{t_i})\,p(\mathcal{A}_{t_{p-1}},\eta,t_p-t_{p-1})\Big].&\end{aligned}$$

The two latter results, together with (4.13), (4.14), the finiteness of the sums and the Markov property applied to $\mathcal{A}$ lead to

$$\begin{aligned}\lim_{D\to\infty} \mathbf{E}_\zeta\Big[\prod_{i=1}^{p} f_i(\mathcal{A}^D_{t_i})\Big] &= \sum_{\eta\in \mathrm{P}^s_n} f_p(\eta)\, \mathbb{E}_{\underline{\zeta}}\Big[\prod_{i=1}^{p-1} f_i(\mathcal{A}_{t_i})\, p(\mathcal{A}_{t_{p-1}},\eta, t_p - t_{p-1})\Big] \\ &= \mathbb{E}_{\underline{\zeta}}\Big[\prod_{i=1}^{p} f_i(\mathcal{A}_{t_i})\Big].\end{aligned}$$

As any real-valued function on $\big(\mathrm{P}^s_n\big)^p$ can be obtained as a uniform limit of product functions, the convergence of the p-dimensional distributions is proven. The proof of Theorem 4.4 is complete by the induction principle. ☐

Proof of Lemma 4.20. Let us start by proving that σ^D_1 converges in probability to 0. If $\zeta \in \Pi_n$, then $\sigma^D_1 = 0$ a.s. for all D so the convergence trivially holds. If $\zeta \notin \Pi_n$, then $\sigma^D_1 > 0$ a.s. and with the notation introduced previously, we have for each function f on P^s_n

$$G^D f(\zeta) = r_D \Psi f(\zeta) + \Gamma f(\zeta) + R_D f(\zeta),$$

where $\Psi f(\zeta) \neq 0$ (in fact, this holds for any $\eta \notin \Pi_n$, and consequently for all values of $\mathcal{A}^D_t,\ t \in [0,\sigma^D_1)$). Let us write $r_D c_\Psi(\zeta)$ (resp. $c_\Gamma(\zeta)$, $c_{R_D}(\zeta)$) the total rate of the non-trivial events generated by $r_D\Psi$ (resp. Γ, R_D) when $G^D f$ is applied to ζ. As events are discrete for each D, we can write for $s > 0$

$$\begin{aligned}\mathbf{P}_\zeta\big[\sigma^D_1 > s\big] \le\, &\mathbf{P}_\zeta\big[\text{at most } n\ \Psi\text{−events and then a } \Gamma\text{− or } R_D\text{−event occur in } [0,s] \\ &\qquad \text{and } \mathcal{A}^D_u \notin \Pi_n\ \forall u \in [0,s]\big] \\ &+\mathbf{P}_\zeta\big[\text{at most } n\ \Psi\text{−events and no } \Gamma\text{− or } R_D\text{−events occur in } [0,s] \\ &\qquad \text{and } \mathcal{A}^D_u \notin \Pi_n\ \forall u \in [0,s]\big] \\ &+\mathbf{P}_\zeta\big[\text{more than } n\ \Psi\text{−events occur before the first } \Gamma\text{− or } R_D\text{−event}\big].\end{aligned} \tag{4.17}$$

Since the events generated by Ψ correspond to the structured genealogical process $(\xi_t, t \ge 0)$ started at ζ as long as no Γ- or R_D-events occurred, by the bound on the number of transitions of ξ (n, see the previous section), the third term on the right-hand side of (4.17) vanishes. Moreover, the probability that the next event generated by G^D is a Γ or an R_D-event when the current value of $\mathcal{A}^D$ is $\eta \notin \Pi_n$ is given by

$$\frac{c_\Gamma(\eta) + c_{R_D}(\eta)}{c_\Gamma(\eta) + c_{R_D}(\eta) + r_D c_\Psi(\eta)} \to 0, \qquad D \to \infty,$$

since $c_\Psi(\eta) > 0$ for such an η, and this is precisely the kind of situation required to be in the configuration given by the first term of (4.17). So by bounding this term by the maximum over $\eta \notin \Pi_n$ of the probabilities calculated just before, we obtain that the first term of (4.17) tends to 0 as D grows to infinity. To finish, for each D and all $k \in \{1,\dots,n\}$ let us call U^D_k the random time of the k'th event occurring to $\mathcal{A}^D$, with the convention that $U^D_k = +\infty$ if there are less than k such events. If k events occur (i.e., $U^D_k < \infty$) and $\mathcal{A}^D$ stays out of Π_n, then U_{k+1} is stochastically bounded by the sum of $k+1$ i.i.d. exponential variables with parameter $r_D \min_{\eta\notin\Pi_n} c_\Psi(\eta)$, whose distribution becomes concentrated close to 0 as D grows since $\min_{\eta\notin\Pi_n} c_\Psi(\eta) > 0$. Consequently,

$$\begin{aligned}&\mathbf{P}_\zeta\big[\text{exactly } k\ \Psi\text{−events and no } \Gamma\text{− or } R_D\text{−events occur in } [0,s],\ \text{and } \mathcal{A}^D_u \notin \Pi_n\ \forall u \in [0,s]\big] \\ &\qquad \le \mathbf{P}_\zeta\big[U^D_{k+1} > s,\ U^D_k < \infty \text{ and } \mathcal{A}^D_u \notin \Pi_n\ \forall u \in [0,s]\big] \to 0.\end{aligned}$$

As the second term in (4.17) is bounded by the sum over $k \in \{0, \dots, n\}$ of the preceding quantities, it converges to zero. Hence, $\mathbf{P}_\zeta[\sigma_1^D > s] \to 0$ for all $s > 0$ and $\sigma_1 \to 0$ in probability.

Now, let f be a function on P_n^s. For each $s > 0$, we have

$$\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D})\,\mathbb{I}_{\{\sigma_1^D < s\}}\big] = \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D})\big] - \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D})\,\mathbb{I}_{\{\sigma_1^D \geq s\}}\big]. \tag{4.18}$$

By the convergence in probability of σ_1^D to 0 and the fact that f is bounded, the second term in the right-hand side of (4.18) vanishes as D grows to infinity. Furthermore, we have

$$\begin{aligned} \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D})\big] = {} & \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D});\ \text{only } \Psi\text{−events before } \sigma_1^D\big] \\ & + \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D});\ \text{at least one } \Gamma\text{−or } R_D\text{−events before } \sigma_1^D\big]. \end{aligned} \tag{4.19}$$

The second term in (4.19) is bounded by $\|f\|\ \mathbf{P}_\zeta\big[\text{at least one } \Gamma\text{−or } R_D\text{−events before } \sigma_1^D\big]$, which tends to 0 by the preceding calculations. Notice that this bound gives us as a by-product that $\mathbf{P}_\zeta\big[\text{only } \Psi\text{−events before } \sigma_1^D\big]$ is positive for D large enough. Moreover, when only $\Psi-$events occurred between 0 and σ_1^D, then the evolution of $\mathcal{A}^D$ is driven by the structured genealogical process ξ started at ζ, so $\mathcal{A}^D_{\sigma_1^D}$ has the same distribution as $\underline{\zeta}$. Thus,

$$\begin{aligned} & \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D})\,;\ \text{only } \Psi\text{−events before } \sigma_1^D\big] \\ & \qquad = \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D})\big|\ \text{only } \Psi\text{−events before } \sigma_1^D\big]\mathbf{P}_\zeta\big[\text{only } \Psi\text{−events before } \sigma_1^D\big] \\ & \qquad = \mathbf{E}\big[f(\underline{\zeta})\big]\mathbf{P}_\zeta\big[\text{only } \Psi\text{−events before } \sigma_1^D\big] \\ & \qquad \to \mathbf{E}\big[f(\underline{\zeta})\big]. \end{aligned}$$

Together with (4.18) and (4.19), we obtain that

$$\lim_{D\to\infty} \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_1^D})\,\mathbb{I}_{\{\sigma_1^D < s\}}\big] = \mathbf{E}[f(\underline{\zeta})] = \mathbb{E}_{\underline{\zeta}}[f(\mathcal{A}_0)\,\mathbb{I}_{\{0<s\}}].$$

A monotone class argument of enables us to conclude the same result for any bounded measurable function f on $\mathbf{R}_+ \times \mathrm{P}_n^s$.

Let us now investigate the convergence of τ_1^D. Recall that if $\eta \in \Pi_n$ and $f \in \mathcal{D}(G^D)$, then

$$G^D f(\eta) = \Gamma f(\eta) + R_D f(\eta).$$

If $s > 0$, by the strong Markov property applied to $\mathcal{A}^D$ at time σ_1^D we have

$$\mathbf{P}_\zeta[\tau_1^D > s] = \mathbf{E}_\zeta\Big[\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}[\tilde{\tau}_1^D > s - \sigma_1^D]\,\mathbb{I}_{\{s > \sigma_1^D\}}\Big] + \mathbf{E}_\zeta\Big[\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}[\tilde{\tau}_1^D > s - \sigma_1^D]\,\mathbb{I}_{\{s \leq \sigma_1^D < \infty\}}\Big]. \tag{4.20}$$

The second term in (4.20) is equal to $\mathbf{P}_\zeta[s \leq \sigma_1^D < \infty]$ which tends to 0 when D grows to infinity. If a Γ-event occurs when the current value of $\mathcal{A}^D$ lies in Π_n, it is necessarily a 2-event, hence the first term is equal to

$$\begin{aligned} & \mathbf{E}_\zeta\ \Big[\ \mathbb{I}_{\{s > \sigma_1^D\}}\ \mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}[\text{no } \Gamma\text{− or } R_D\text{−events between 0 and } s - \sigma_1^D]\Big] \\ & \quad + \mathbf{E}_\zeta\Big[\mathbb{I}_{\{s > \sigma_1^D\}}\ \mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}[\text{no } \Gamma\text{−events and at least one } R_D\text{−event between 0 and } s - \sigma_1^D; \\ & \qquad\qquad \tilde{\tau}_1^D > s - \sigma_1^D]\Big]. \end{aligned}$$

But for all $\eta \in \Pi_n$,

$$\begin{aligned} & \mathbf{P}_\eta\big[\,\text{no } \Gamma\text{−events and at least one } R_D\text{−event between 0 and } s - \sigma_1^D;\ \tilde{\tau}_1^D > s - \sigma_1^D\big] \\ & \quad \leq \mathbf{P}_\eta\big[\text{no } \Gamma\text{−events and at least one } R_D\text{−event between 0 and } s\big] \end{aligned}$$

$$\leq 1-\exp\big(-s\max_{\gamma\in\mathrm{P}_n^s}c_{R_D}(\gamma)\big)\to 0,$$

so by dominated convergence,

$$\begin{aligned}\mathbf{E}_\zeta\Big[\mathbb{I}_{\{s>\sigma_1^D\}}\,\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}&[\text{no }\Gamma-\text{event and at least one }R_D-\text{event between }0\text{ and }s-\sigma_1^D;\\ &\tilde{\tau}_1^D>s-\sigma_1^D]\Big]\to 0.\end{aligned}$$

Consequently,

$$\begin{aligned}\mathbf{P}_\zeta[\tau_1^D>s]&=\mathbf{E}_\zeta\Big[\mathbb{I}_{\{s>\sigma_1^D\}}\,\exp-\Big\{\big(c_\Gamma(\mathcal{A}^D_{\sigma_1^D})+c_{R_D}(\mathcal{A}^D_{\sigma_1^D})\big)(s-\sigma_1^D)\Big\}\Big]\\ &\to\mathbb{E}[e^{-sc_\Gamma(\underline{\zeta})}]=\mathbb{P}_{\underline{\zeta}}[\sigma_2>s]\end{aligned}$$

by the preceding convergence result for $(\sigma_1^D,\mathcal{A}^D_{\sigma_1^D})$ and the uniform convergence of c_{R_D} towards 0. We can thus conclude that the law of τ_1^D under $\mathbf{P}_\zeta$ converges to the law of σ_2 under $\mathbb{P}_{\underline{\zeta}}$.

Now, by the strong Markov property applied to $\mathcal{A}^D$ at time σ_1^D, we have

$$\begin{aligned}&\mathbf{E}_\zeta\Big[\,\mathbb{I}_{\{\tau_1^D<s\}}\,f(\mathcal{A}^D_{\tau_1^D})\,\mathbb{I}_{\{\tau_1^D<\infty\}}\Big]\\ &\quad=\mathbf{E}_\zeta\Big[\,\mathbb{I}_{\{\sigma_1^D<\infty\}}\mathbf{E}_{\mathcal{A}^D_{\sigma_1^D}}[\,\mathbb{I}_{\{\tilde{\tau}_1^D<s-\sigma_1^D\}}\,f(\tilde{\mathcal{A}}^D_{\tilde{\tau}_1^D});\ \text{the first event is an }R_D-\text{event}]\Big]\\ &\quad\quad+\mathbf{E}_\zeta\Big[\,\mathbb{I}_{\{\sigma_1^D<\infty\}}\mathbf{E}_{\mathcal{A}^D_{\sigma_1^D}}[\,\mathbb{I}_{\{\tilde{\tau}_1^D<s-\sigma_1^D\}}\,f(\tilde{\mathcal{A}}^D_{\tilde{\tau}_1^D});\ \text{the first event is a }\Gamma-\text{event}]\Big].\end{aligned}\tag{4.21}$$

The absolute value of the first term in the right-hand side of (4.21) is bounded by

$$\|f\|\max_{\eta\in\Pi_n}\mathbf{P}_\eta[\text{a first event occurs and is an }R_D-\text{event}]\to 0.$$

Moreover, if $\tilde{\mathcal{A}}_0^D=\eta\in\Pi_n$ and the first event is a Γ-event, then $\tilde{\tau}_1^D$ is the time of that first event and $\tilde{\mathcal{A}}^D_{\tilde{\tau}_1^D}$ its outcome. Therefore, both are independent and $\tilde{\mathcal{A}}^D_{\tilde{\tau}_1^D}$ is distributed like $\chi(\eta)$, so the second term in (4.21) is equal to

$$\begin{aligned}&\mathbf{E}_\zeta\Big[\,\mathbb{I}_{\{\sigma_1^D<\infty\}}\mathbf{E}_{\mathcal{A}^D_{\sigma_1^D}}[\,\mathbb{I}_{\{\tilde{\tau}_1^D<s-\sigma_1^D\}}f(\chi(\tilde{\mathcal{A}}_0^D))\,]\Big]+o(1)\\ &\quad=\mathbf{E}_\zeta\Big[\,\mathbb{I}_{\{\sigma_1^D<\infty\}}\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}[\tilde{\tau}_1^D<s-\sigma_1^D]\,\mathbf{E}_{\mathcal{A}^D_{\sigma_1^D}}[f(\chi(\tilde{\mathcal{A}}_0^D))]\,\Big]+o(1).\end{aligned}$$

Let us write

$$\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}\big[\tilde{\tau}_1^D<s-\sigma_1^D\big]=\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}\big[\tilde{\tau}_1^D<s\big]-\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}\big[\tilde{\tau}_1^D\in[s-\sigma_1^D,s]\big]$$

and fix $\epsilon>0$. For any $\delta>0$, we have

$$\begin{aligned}&\mathbf{E}_\zeta\Big[\,\mathbb{I}_{\{\sigma_1^D<\infty\}}\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}\big[\tilde{\tau}_1^D\in[s-\sigma_1^D,s]\big]\,\mathbf{E}_{\mathcal{A}^D_{\sigma_1^D}}[f(\chi(\tilde{\mathcal{A}}_0^D))]\,\Big]\\ &\quad\leq\mathbf{E}_\zeta\Big[\,\mathbb{I}_{\{\sigma_1^D<\delta\}}\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}\big[\tilde{\tau}_1^D\in[s-\delta,s]\big]\,\mathbf{E}_{\mathcal{A}^D_{\sigma_1^D}}[|f(\chi(\tilde{\mathcal{A}}_0^D))|]\,\Big]+\|f\|\,\mathbf{P}_\zeta[\sigma_1^D\in[\delta,\infty)].\end{aligned}$$

By the convergence in probability of σ_1^D to 0, there exists $D_0\geq 1$ such that for all $D\geq D_0$,

$$\mathbf{P}_\zeta\big[\sigma_1^D\in[\delta,\infty)\big]<\frac{\epsilon}{3\|f\|}.$$

Let $\eta\in\Pi_n$. By the continuity of the distribution function of σ_2, there exists $\delta_0>0$ such that

$$\mathbf{P}_\eta\big[\tilde{\sigma}_2\in[s-\delta_0,s]\big]<\frac{\epsilon}{3\|f\|}.$$

In addition, $\tilde{\tau}_1^D$ converges in distribution to σ_2, hence there exists $D_1 \geq 1$ such that for all $D \geq D_1$,

$$\Big|\mathbf{P}_\eta\big[\tilde{\tau}_1^D \in [s-\delta_0, s]\big] - \mathbf{P}_\eta\big[\tilde{\sigma}_2 \in [s-\delta_0, s]\big]\Big| < \frac{\epsilon}{3\|f\|}.$$

Since Π_n is a finite set, we can conclude that for $\delta > 0$ small enough, and D large enough, we have

$$\begin{aligned}
&\mathbf{E}_\zeta\left[\mathbb{I}_{\{\sigma_1^D<\delta\}}\mathbf{P}_{\mathcal{A}^D_{\sigma_1^D}}\big[\tilde{\tau}_1^D \in [s-\delta, s]\big]\,\mathbf{E}_{\mathcal{A}^D_{\sigma_1^D}}[|f(\chi(\tilde{\mathcal{A}}_0^D))|]\right] + \|f\|\,\mathbf{P}_\zeta[\sigma_1^D \in [\delta,\infty)] \\
&\quad \leq \mathbf{E}_\zeta\Big[\,\mathbb{I}_{\{\sigma_1^D<\delta\}}\max_{\eta\in\Pi_n}\mathbf{P}_\eta\big[\tilde{\tau}_1^D \in [s-\delta, s]\big]\,\mathbf{E}_{\mathcal{A}^D_{\sigma_1^D}}[|f(\chi(\tilde{\mathcal{A}}_0^D))|]\,\Big] + \frac{\epsilon}{3} \\
&\quad \leq \epsilon.
\end{aligned}$$

Now, $\eta \mapsto \mathbf{P}_\eta[\tilde{\tau}_1^D < s]$ converges uniformly in $\eta \in \Pi_n$ to $\eta \mapsto \mathbb{P}_\eta[\tilde{\sigma}_2 < s]$ and

$$(s,\eta) \mapsto \mathbb{I}_{\{s<\infty\}}\mathbb{I}_{\{\eta\in\Pi_n\}}\mathbb{P}_\eta[\tilde{\sigma}_2 < s]\mathbb{E}_\eta\big[f(\chi(\tilde{\mathcal{A}}))\big]$$

is a bounded measurable function, so by the convergence in distribution of $(\sigma_1^D, \mathcal{A}^D_{\sigma_1^D})$ proven above, for D large enough we have

$$\begin{aligned}
&\Big|\mathbf{E}_\zeta\left[\mathbb{I}_{\{\tau_1^D<s\}}\,f(\mathcal{A}^D_{\tau_1^D})\,\mathbb{I}_{\{\tau_1^D<\infty\}}\right] - \mathbb{E}_{\underline{\zeta}}\big[\mathbb{I}_{\{\sigma_2<s\}}\,f(\chi(\mathcal{A}_{\sigma_1}))\,\mathbb{I}_{\{\sigma_2<\infty\}}\big]\Big| \\
&\quad = \Big|\mathbf{E}_\zeta\big[\mathbb{I}_{\{\tau_1^D<s\}}\,f(\mathcal{A}^D_{\tau_1^D})\,\mathbb{I}_{\{\tau_1^D<\infty\}}\big] - \mathbb{E}_{\underline{\zeta}}\big[\mathbb{I}_{\{\sigma_1<\infty\}}\mathbb{P}_{\mathcal{A}_{\sigma_1}}[\tilde{\sigma}_2 < s]\,\mathbb{E}_{\mathcal{A}_{\sigma_1}}\big[f(\chi(\tilde{\mathcal{A}}_0))\big]\big]\Big| \\
&\quad < 3\epsilon.
\end{aligned}$$

Letting ϵ tend to zero yields the desired result (once again by invoking monotone classes) and completes the step $i = 1$ of the proof of Lemma 4.20.

Suppose that the desired properties hold for $i-1$. Let f be a bounded continuous function on $\mathbf{R}_+ \times \mathrm{P}_n^s$. Since $\mathbb{I}_{\{\sigma_i^D<\infty\}} = \mathbb{I}_{\{\sigma_i^D<\infty\}}\mathbb{I}_{\{\tau_{i-1}^D<\infty\}}$, the strong Markov property applied to $\mathcal{A}^D$ at time τ_{i-1}^D gives

$$\mathbf{E}_\zeta\Big[f(\sigma_i^D, \mathcal{A}^D_{\sigma_i^D})\mathbb{I}_{\{\sigma_i^D<\infty\}}\Big] = \mathbf{E}_\zeta\Big[\mathbb{I}_{\{\tau_{i-1}^D<\infty\}}\mathbf{E}_{\mathcal{A}^D_{\tau_{i-1}^D}}\,\big[f(\tau_{i-1}^D + \tilde{\sigma}_1^D, \tilde{\mathcal{A}}^D_{\tilde{\sigma}_1^D})\,\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\big]\Big]$$

But, for all $\eta \notin \Pi_n$, if X denotes a random variable whose distribution under $\mathbf{P}_\eta$ is that of $\underline{\eta}$ (e.g. ξ_{τ_π} in the notation of Proposition 4.10), then

$$\begin{aligned}
&\Big|\mathbf{E}_\eta\Big[f(t+\tilde{\sigma}_1^D, \tilde{\mathcal{A}}^D_{\tilde{\sigma}_1^D})\,\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\Big] - \mathbf{E}_\eta[f(t,X)]\Big| \\
&\qquad \leq \Big|\mathbf{E}_\eta\Big[f(t+\tilde{\sigma}_1^D, \tilde{\mathcal{A}}^D_{\tilde{\sigma}_1^D})\,\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\Big] - \mathbf{E}_\eta\big[f(t+\tilde{\sigma}_1^D, X)\,\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\big]\Big| \\
&\qquad\quad + \Big|\mathbf{E}_\eta\big[f(t+\tilde{\sigma}_1^D, X)\,\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\big] - \mathbf{E}_\eta[f(t,X)]\Big|.
\end{aligned}$$

Since $\tilde{\mathcal{A}}^D_{\tilde{\sigma}_1^D}$ has the same distribution as $\underline{\eta}$ under $\mathbf{P}_\eta$ if only Ψ-events occurred between 0 and $\tilde{\sigma}_1^D$, the first term is equal to

$$\begin{aligned}
&\Big|\mathbf{E}_\eta\left[\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\big(f(t+\tilde{\sigma}_1^D, \tilde{\mathcal{A}}^D_{\tilde{\sigma}_1^D}) - f(t+\tilde{\sigma}_1^D, X)\big);\ \text{not only } \Psi\text{−events between 0 and } \tilde{\sigma}_1^D\right]\Big| \\
&\qquad \leq 2\|f\|\max_{\eta\notin\Pi_n}\mathbf{P}_\eta\big[\tilde{\sigma}_1^D < \infty, \text{not only } \Psi\text{−events between 0 and } \tilde{\sigma}_1^D\big] \to 0
\end{aligned}$$

by the calculations done in the proof of the convergence of σ_1^D. Moreover,

$$\mathbf{E}_\eta\big[f(t+\tilde{\sigma}_1^D, X)\,\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\big] - \mathbf{E}_\eta[f(t,X)] \to 0$$

uniformly in η by the convergence in probability of $\tilde{\sigma}_1^D$ towards 0 and the finiteness of the number of states that X can take. Therefore,

$$\mathbf{E}_\eta\Big[f(t+\tilde{\sigma}_1^D,\tilde{\mathcal{A}}^D_{\tilde{\sigma}_1^D})\,\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\Big] \to \mathbf{E}_\eta[f(t,X)] = \mathbb{E}_{\underline{\eta}}[f(t,\mathcal{A}_0)]$$

uniformly in (t,η). This uniform convergence (which trivially holds also for $\eta \in \Pi_n$ since $\underline{\eta} = \eta$ and $\sigma_1^D = 0$ a.s.), together with the induction hypothesis for $i-1$ yields

$$\lim_{D\to\infty}\mathbf{E}_\zeta\Big[\mathbb{I}_{\{\tau_{i-1}^D<\infty\}}\mathbf{E}_{\mathcal{A}^D_{\tau_{i-1}^D}}\big[f(\tau_{i-1}^D+\tilde{\sigma}_1^D,\tilde{\mathcal{A}}^D_{\tilde{\sigma}_1^D})\,\mathbb{I}_{\{\tilde{\sigma}_1^D<\infty\}}\big]\Big] = \mathbb{E}_{\underline{\zeta}}\big[\mathbb{I}_{\{\sigma_i<\infty\}}\mathbb{E}_{\chi(\mathcal{A}_{\sigma_{i-1}})}[f(\sigma_i,X)]\big].$$

But, from the description of the evolution of $\mathcal{A}$ in terms of the geographical events followed by the instantaneous action of the structured genealogical process ξ, we see that the random variable $\underline{\chi(\mathcal{A}_{\sigma_{i-1}})}$ is distributed precisely like $\mathcal{A}_{\sigma_i}$ (if $\sigma_i < \infty$). Consequently,

$$\lim_{D\to\infty}\mathbf{E}_\zeta\Big[f(\sigma_i^D,\mathcal{A}^D_{\sigma_i^D})\mathbb{I}_{\{\sigma_i^D<\infty\}}\Big] = \mathbb{E}_{\underline{\zeta}}\big[\mathbb{I}_{\{\sigma_i<\infty\}}f(\sigma_i,\mathcal{A}_{\sigma_i})\big].$$

The same technique applies to $(\tau_i^D,\mathcal{A}^D_{\tau_i^D})$, where this time we use the strong Markov property at time σ_i^D and the following convergence result:

$$\mathbf{E}_\eta\Big[f(t+\tau_1^D,\mathcal{A}^D_{\tau_1^D})\,\mathbb{I}_{\{\tau_1^D<\infty\}}\Big] \to \mathbb{E}_\eta\big[f(t+\sigma_2,\chi(Y))\,\mathbb{I}_{\{\sigma_2<\infty\}}\big]$$

uniformly in $(t,\eta) \in \mathbf{R}_+ \times \Pi_n$, where under $\mathbf{P}_\eta$, Y is a.s. equal to η. □

Proof of Lemma 4.21. If an event occurs in the (random) interval $[\sigma_i^D,\tau_i^D)$, the first such event can be neither a Ψ-event since $\mathcal{A}^D_{\sigma_i^D} \in \Pi_n$, nor a Γ-event since $\mathcal{A}^D$ would then undergo a 2-event before time τ_i^D, contradicting the definition of τ_i^D, so it must be an R_D-event. Consequently, if we write

$$\begin{aligned}
&\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\;\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\big|\sigma_i^D<\infty\big]\\
&\qquad= \mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\;\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t);\ \text{nothing happens in } [\sigma_i^D,t]\big|\ \sigma_i^D<\infty\big]\\
&\qquad\quad+\mathbf{E}_\zeta\big[f(\mathcal{A}_t^D)\;\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t);\ \text{something occurs in } [\sigma_i^D,t]\big|\ \sigma_i^D<\infty\big],
\end{aligned}\tag{4.22}$$

then the absolute value of the second term of the right-hand side of (4.22) is bounded by

$$\begin{aligned}
&\|f\|\;\mathbf{P}_\zeta\big[\sigma_i^D \le t \text{ and an } R_D\text{−event occurs in } [\sigma_i^D,t]\big|\ \sigma_i^D<\infty\big]\\
&\qquad\le \|f\|\big(1-\exp(-t\max_\eta c_{R_D}(\eta))\big) \to 0,
\end{aligned}$$

where in the exponential the maximum is over $\eta \in \mathrm{P}_n^s$ and recall that $c_{R_D}(\eta)$ is the total rate at which R_D-events occur when the current value of $\mathcal{A}^D$ is η.

The first term of (4.22) is equal to

$$\begin{aligned}
&\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\;\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\ ;\ \text{nothing happens in } [\sigma_i^D,t]\big|\ \sigma_i^D<\infty\big]\\
&\qquad= \mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\;\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\ \big|\ \sigma_i^D<\infty\big]\\
&\qquad\quad-\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\;\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\ ;\ \text{something happens in } [\sigma_i^D,t]\ \big|\ \sigma_i^D<\infty\big].
\end{aligned}$$

As before,

$$\begin{aligned}
&\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\;\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t);\ \text{something happens in } [\sigma_i^D,t]\big|\ \sigma_i^D<\infty\big]\\
&\qquad\le \|f\|\big(1-\exp(-t\max_\eta c_{R_D}(\eta))\big) \to 0
\end{aligned}$$

and furthermore

$$\begin{aligned}\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\,\mathbb{I}_{[\sigma_i^D,\tau_i^D)}(t)\mid \sigma_i^D<\infty\big] = {}&\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\,\mathbb{I}_{\{\sigma_i^D\le t\}}\mid \sigma_i^D<\infty\big]\\ &-\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\,\mathbb{I}_{\{\tau_i^D\le t\}}\mid \sigma_i^D<\infty\big].\end{aligned} \tag{4.23}$$

On the one hand, by Lemma 4.20 and the fact that $\mathbf{P}_\zeta[\sigma_i^D<\infty]\to\mathbb{P}_{\underline{\zeta}}[\sigma_i<\infty]>0$, the first term in (4.23) converges as D tends to infinity to

$$\mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_{\sigma_i})\,\mathbb{I}_{\{\sigma_i\le t\}}\mid \sigma_i<\infty\big].$$

On the other hand, by the strong Markov property applied to $\mathcal{A}^D$ at time σ_i^D, the second term in (4.23) is equal to

$$\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\,\mathbf{P}_{\mathcal{A}^D_{\sigma_i^D}}\big[\tilde{\tau}_1^D\le t\big]\mid \sigma_i^D<\infty\big].$$

The function $\eta\mapsto\mathbf{P}_\eta\big[\tilde{\tau}_1^D\le t\big]$ converges uniformly in $\eta\in\Pi_n$ to $\eta\mapsto\mathbb{P}_\eta\big[\tilde{\sigma}_2\le t\big]$, so by Lemma 4.20 we obtain

$$\begin{aligned}\lim_{D\to\infty}\mathbf{E}_\zeta\big[f(\mathcal{A}^D_{\sigma_i^D})\,\mathbf{P}_{\mathcal{A}^D_{\sigma_i^D}}\big[\tilde{\tau}_1^D\le t\big]\mid \sigma_i^D<\infty\big] &= \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_{\sigma_i})\,\mathbb{P}_{\mathcal{A}_{\sigma_i}}[\tilde{\sigma}_2\le t]\mid \sigma_i<\infty\big]\\ &= \mathbb{E}_{\underline{\zeta}}\big[f(\mathcal{A}_{\sigma_i})\,\mathbb{I}_{\{\sigma_{i+1}\le t\}}\mid \sigma_i<\infty\big]\end{aligned}$$

by the strong Markov property applied this time to $\mathcal{A}$ at time σ_i. Combining the above, we obtain the desired result. □

The results obtained in this section are similar in spirit to perturbation theorems such as Theorem 1.7.6 in [**EK86**]. Indeed, in our case the existence of a projector p corresponding to Ψ and the convergence of the semigroup required (see Condition (7.12) and Remark 1.7.5 in [**EK86**], p.39) easily follows from Lemma 4.10 and the finiteness of P_n^s. Furthermore, the existence of a limit for $r_D^{-1}G^D$ is obvious from the form of G^D. However, Condition (7.17) of Theorem 1.7.6 requires the existence of a subspace E of functions on P_n^s such that for every $f\in E$, there exist functions $g, f_1, f_2, \ldots$ satisfying

$$\|f-f_D\|\to 0 \qquad \text{and} \qquad \|g-G^Df_D\|\to 0 \qquad \text{as } D\to\infty.$$

The condition on $(f_D)_{D\ge 1}$ and the finiteness of P_n^s yield

$$G^Df_D = r_D\Psi f + o(r_D),$$

implying that a corresponding function g can exist only if $\Psi f\equiv 0$. Although $\Psi f(\zeta)=0$ if $\zeta\in\Pi_n$, this condition would also require that $f(\zeta)=0$ whenever $\zeta\notin\Pi_n$. Hence, to fit into Ethier and Kurtz' framework, an obvious candidate for E would be

$$E=\{f : f(\zeta)=0 \text{ for all } \zeta\in\mathrm{P}_n^s\setminus\Pi_n\},$$

where we then define a bounded linear transformation $\wp_n : \mathrm{P}_n^s\to\Pi_n$ such that $\wp_n(\eta)=\eta$ for every $\eta\in\Pi_n$. We may then apply Theorem 1.7.6 of [**EK86**] and obtain convergence of the semigroup corresponding to (or equivalently here of the finite dimensional distributions of) $\wp_n(\mathcal{A}^D)$ to that of $\mathcal{A}$. However, it is unclear how to define $\wp_n$ on the set $\mathrm{P}_n^s\setminus\Pi_n$, that is to specify how to project P_n^s down onto its subset Π_n, in such a way that the operator $\big\{\big(f\circ\wp_n, (\Gamma^n\circ p)(f\circ\wp_n)\big),\ f:\mathrm{P}_n^s\to\mathrm{P}_n^s\big\}$ generates a Markov process. Unfortunately, unless this condition is satisfied, Theorem 1.7.6 cannot be used to deduce the convergence result given in our Theorem 4.4.

3.3. Tightness. The convergence of the finite-dimensional distributions relies on the fact that the time required for the process to re-enter Π_n following a geographical collision is vanishingly small as D tends to infinity. On the other hand, multiple changes to the configuration of the genealogy can occur during this short period with high probability, so that the conditions for $\mathcal{A}^D$ to converge as processes in $D_{\mathrm{P}_n^s}[0,\infty)$ are much more delicate.

Recall the definition of the stopping times σ_i^D and τ_i^D given in (4.8). Suppose that the probability that a 2-event results in a configuration not in Π_n vanishes as $D \to 0$, or equivalently that

$$\lim_{D\to\infty} \mathbf{P}_\zeta\big[\tau_1^D < \infty,\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_2\big] = 0, \tag{4.24}$$

where $\zeta = (\{\{1\}\},\{\{2\}\})$ and the equivalence stems from the consistency equations (4.4). Then, we easily see that the first time τ after σ_1^D such that $\mathcal{A}^D \notin \Pi_n$ converges to $+\infty$ in probability. Since

$$G^D f(\eta) = \Gamma f(\eta) + o(1) \qquad \text{as } D \to \infty$$

if $\eta \in \Pi_n$, we readily obtain that for any $a > 0$, the sequence of processes $(\{\mathcal{A}_t^D, t \geq a\},\ D \geq 1)$ is tight (recall that σ_1^D converges in probability towards 0). Let us now show that if Condition (4.24) does not hold, the sequence $\mathcal{A}^D$ is not tight. It will be easier to work with a metric on P_n^s, the associated topology still being the discrete topology.

PROPOSITION 4.22. *Assume that (4.24) does not hold. Let d be a discrete metric on* P_n^s*, and suppose that* $\zeta \in \mathrm{P}_n^s$ *is such that* $\mathbb{P}_{\underline{\zeta}}[\sigma_2 < \infty] > 0$. *Then the sequence of processes* $\mathcal{A}^D$ *under* $\mathbf{P}_\zeta$ *is not tight in* $D_{\mathrm{P}_n^s}[0,\infty)$ *endowed with the Skorokhod topology corresponding to d.*

Proof. First, recall the definition of the modulus of continuity w' given in [**EK86**], p.122: for $X \in D_{\mathrm{P}_n^s}[0,\infty)$, $\delta > 0$ and $T > 0$,

$$w'(X,\delta,T) := \inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} d(X_s,X_t), \tag{4.25}$$

where the infimum is over all finite sets of times of the form $0 = t_0 < t_1 < \cdots < t_{n-1} < T \leq t_n$ such that $\min_{1\leq i\leq n}(t_i - t_{i-1}) > \delta$ and $n \geq 1$.

Suppose that the sequence $\mathcal{A}^D$ is tight. P_n^s is a finite set, so the discrete topology on (P_n^s, d) turns it into a complete and separable metric space, therefore $\mathcal{A}^D$ is also relatively compact. By Corollary 3.7.4 of [**EK86**], this implies that for every $\gamma \in \mathrm{P}_n^s$, all $\eta > 0$ and $T > 0$, there exists $\delta > 0$ such that

$$\limsup_{D\to\infty} \mathbf{P}_\gamma\big[w'(\mathcal{A}^D,\delta,T) \geq \eta\big] \leq \eta. \tag{4.26}$$

Besides, the finiteness of P_n^s guarantees the existence of $\epsilon > 0$ such that, if $\gamma \neq \gamma' \in \mathrm{P}_n^s$, then $d(\gamma,\gamma') > \epsilon$.

Let $T = 1$, $\eta \in (0,\epsilon)$ and $\delta \in (0,1)$. We have

$$\begin{aligned}\mathbf{P}_\zeta\big[w'(\mathcal{A}^D,\delta,1) \geq \eta\big] = {}& \mathbf{P}_\zeta\Big[w'(\mathcal{A}^D,\delta,1) \geq \eta \,\Big|\, \tau_1^D \geq \frac{1}{2}\Big]\mathbf{P}_\zeta\Big[\tau_1^D \geq \frac{1}{2}\Big] \\ & + \mathbf{P}_\zeta\Big[w'(\mathcal{A}^D,\delta,1) \geq \eta \,\Big|\, \tau_1^D < \frac{1}{2}\Big]\mathbf{P}_\zeta\Big[\tau_1^D < \frac{1}{2}\Big].\end{aligned} \tag{4.27}$$

On the one hand, σ_1^D converges to 0 in probability and $\tau_1^D \Rightarrow \sigma_2$, so by Slutsky's lemma (see Lemma 2.8 in [**vdV98**]) $\tau_1^D - \sigma_1^D \Rightarrow \sigma_2$, which is an exponential random variable with positive parameter, so we have

$$\begin{aligned}\mathbf{P}_\zeta\Big[\tau_1^D < \frac{1}{2}\Big] &\geq \mathbf{P}_\zeta\Big[\sigma_1^D < \frac{1}{3} \text{ and } \tau_1^D - \sigma_1^D \leq \frac{1}{6}\Big] \\ &= \mathbf{P}_\zeta\Big[\tau_1^D - \sigma_1^D \leq \frac{1}{6}\Big] - \mathbf{P}_\zeta\Big[\sigma_1^D \geq \frac{1}{3} \text{ and } \tau_1^D - \sigma_1^D \leq \frac{1}{6}\Big]\end{aligned}$$

$$\to \mathbb{P}_{\underline{\zeta}}\Big[\sigma_2 \leq \frac{1}{6}\Big] =: C > 0$$

since the last term in the second line vanishes by the convergence in probability of σ_1^D to 0. On the other hand,

$$\mathbf{P}_\zeta\Big[w'(\mathcal{A}^D,\delta,1) \geq \eta \,\Big|\, \tau_1^D < \frac{1}{2}\Big] \geq \mathbf{P}_\zeta\Big[w'(\mathcal{A}^D,\delta,1) \geq \eta \,\Big|\, \tau_1^D < \frac{1}{2},\ \sigma_2^D - \tau_1^D < \frac{\delta}{2},\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_n\Big]$$
$$\times \mathbf{P}_\zeta\Big[\sigma_2^D - \tau_1^D < \frac{\delta}{2},\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_n \Big|\, \tau_1^D < \frac{1}{2}\Big]. \tag{4.28}$$

By the convergence in probability of σ_1^D to 0, uniformly in $\eta \in \mathrm{P}_n^s$, and the strong Markov property applied to $\mathcal{A}^D$ at time τ_1^D, we obtain that $\sigma_2^D - \tau_1^D$ converges in probability to 0. Furthermore, on the event that no R_D-events occurred between the times σ_1^D and τ_1^D (the probability of which is growing to one), τ_1^D is the epoch of the first event after σ_1^D and $\mathcal{A}^D_{\tau_1^D}$ its outcome so, by the strong Markov property, τ_1^D and $\mathcal{A}^D_{\tau_1^D}$ are independent conditionally on $\mathcal{F}^D_{\sigma_1^D}$. Since (4.24) does not hold, we can write

$$\liminf_{D\to\infty} \mathbf{P}_\zeta\Big[\sigma_2^D - \tau_1^D < \frac{\delta}{2},\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_n \Big|\, \tau_1^D < \frac{1}{2}\Big] \geq \liminf_{D\to\infty} \mathbf{P}_\zeta\big[\mathcal{A}^D_{\tau_1^D} \notin \Pi_n\big] > 0.$$

Now, if $\tau_1^D < \frac{1}{2}$, $\mathcal{A}^D_{\tau_1^D} \notin \Pi_n$ and $\sigma_2^D - \tau_1^D < \frac{\delta}{2}$, then by definition of ϵ,

$$d\big(\mathcal{A}^D_{\tau_1^D-}, \mathcal{A}^D_{\tau_1^D}\big) > \epsilon \quad \text{and} \quad d\big(\mathcal{A}^D_{\tau_1^D}, \mathcal{A}^D_{\sigma_2^D}\big) > \epsilon,$$

and by assumption $\sigma_2^D - \tau_1^D < \frac{\delta}{2}$, so $w'(\mathcal{A}^D,\delta,1) \geq \epsilon \geq \eta$. Consequently,

$$\liminf_{D\to\infty} \mathbf{P}_\zeta\Big[w'(\mathcal{A}^D,\delta,1) \geq \eta \,\Big|\, \tau_1^D < \frac{1}{2}\Big] > 0.$$

Therefore, we see from (4.27) that, for all $\delta \in (0,1)$,

$$\liminf_{D\to\infty} \mathbf{P}_\zeta\big[w'(\mathcal{A}^D,\delta,1) \geq \eta\big] > C' > \eta$$

for any $\eta \in (0, \epsilon \wedge C')$. This yields a contradiction with (4.26). ☐

From the last proof, we see that what prevents the sequence of structured genealogical processes from being tight is that at each geographical collision, at least two jumps accumulate: the geographical collision itself and one or more transitions generated by ξ to bring $\mathcal{A}^D$ back into Π_n.

Yet the unstructured genealogical process, which is not a Markov process for $D < \infty$, is not modified by movements of blocks. Thus, if the number of jumps needed by $\mathcal{A}^D$ to re-enter Π_n after a geographical collision were at most one with a probability growing to 1, we would expect tightness to hold for $\mathcal{A}^{D,u}$ (recall that ζ^u denotes the unstructured partition generated by ζ). The next proposition in fact gives an equivalence between the behaviour of the latter probability and tightness of $\{\mathcal{A}^{D,u}, D \geq 1\}$.

PROPOSITION 4.23. *For each $D \in \mathbb{N}$, let U_1^D be the random time defined by*

$$U_1^D := \inf\big\{t > 0 :\ \mathcal{A}^D_{t-} \neq \mathcal{A}^D_t\big\},$$

with the convention that $\inf\emptyset = +\infty$. *Note that, if $\mathcal{A}_0^D \notin \Pi_n$, then $U_1^D \leq \sigma_1^D$. Let also $\chi(\Pi_n)$ denote the image of Π_n by the first geographical collision (when it exists), that is*

$$\chi(\Pi_n) := \big\{\gamma \in \mathrm{P}_n^s\ :\ \exists\, \zeta \in \Pi_n,\ \mathbb{P}[\chi(\zeta) = \gamma] > 0\big\}.$$

Suppose that for all $\gamma \notin \Pi_n$, $\mathbf{P}[\underline{\gamma}^u \neq \gamma^u] > 0$ (meaning that the process ξ started at γ has at least one coalescence with positive probability).

Then the following are equivalent:

(i) For all $\gamma \in \chi(\Pi_n) \setminus \Pi_n$*,* $\lim_{D\to\infty} \mathbf{P}_\gamma[U_1^D = \sigma_1^D] = 1$.

(ii) For all $\zeta \in \mathrm{P}_n^s$ *and* $a > 0$*, the sequence of* $D_{\mathcal{P}_n}[a,\infty)$*-valued random variables* $\mathcal{A}^{D,u}$*, started at* ζ^u *at time* 0*, is tight.*

Furthermore, if $\zeta \in \Pi_n \cup \chi(\Pi_n)$*, then Condition (i) is equivalent to the tightness in* $D_{\mathcal{P}_n}[0,\infty)$ *of* $\mathcal{A}^{D,u}$ *started at* ζ^u*.*

As a consequence of Theorem 4.4, if Conditions (i) and (ii) hold, then for all $\zeta \in \mathrm{P}_n^s$ *and* $a > 0$*, the law of* $(\mathcal{A}_t^{D,u},\ t \geq a)$ *under* $\mathbf{P}_\zeta$ *converges to the law under* $\mathbb{P}_{\underline{\zeta}^u}$ *of* $(\mathcal{A}_t^u,\ t \geq a)$*. Furthermore, if* $\zeta \in \Pi_n$*, then the convergence holds for* $a = 0$*.*

REMARK 4.24. *Assuming that for all* $\gamma \notin \Pi_n$*,* $\mathbf{P}[\underline{\gamma}^u \neq \gamma^u] > 0$ *is actually not required, but not supposing it makes the proof unnecessarily more involved.*

Proof. Once again we work with a metric d on $\mathcal{P}_n$, so that $D_{\mathcal{P}_n}[0,\infty)$ is a complete and separable metric space and the sequence $(\mathcal{A}^{D,u})_{D\geq 1}$ is tight if and only if it is relatively compact. We call $\epsilon > 0$ the minimum distance between two different partitions. Let us first show that if Condition (i) is not fulfilled, then neither is Condition (ii). The following proof is highly reminescent to the proof of Proposition 4.22, so let us adopt directly the same notation. In particular, we work with $T = 1$ and ζ such that $\mathbb{P}_{\underline{\zeta}}[\sigma_2 < \infty] > 0$.

For each $a > 0$, let us write w'_a the modulus of continuity of a process corresponding to times $t \geq a$, defined as in (4.25) with the condition on the finite sets $\{t_i\}$ replaced by $a = t_0 < \cdots < T \leq t_n$. Fix $a \in (0, 1/3)$, and let $\eta \in (0,\epsilon)$ and $\delta \in (0,1)$. The same calculation as in the proof of Proposition 4.22 holds by replacing the event $\{w'(\mathcal{A}^D,\delta,1) \geq \eta\}$ by $\{w'_a(\mathcal{A}^{D,u},\delta,1) \geq \eta\}$ and $\mathbf{P}_\zeta[\tau_1^D < 1/2]$ by $\mathbf{P}_\zeta[1/3 \leq \tau_1^D < 1/2]$. Hence, by (4.28) and the argument directly following it, we just need to prove that

$$\mathbf{P}_\zeta\Big[w'_a(\mathcal{A}^{D,u},\delta,1) \geq \eta \;\Big|\; \frac{1}{3} \leq \tau_1^D < \frac{1}{2},\ \sigma_2^D - \tau_1^D < \frac{\delta}{2},\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_n\Big] \tag{4.29}$$

is bounded below by a positive constant for D large enough. If we define V_1^D by

$$V_1^D := \inf\big\{t > \tau_1^D \;:\; \mathcal{A}^D_{t-} \neq \mathcal{A}^D_t\big\},$$

then the expression in (4.29) is equal to

$$\begin{aligned}\mathbf{P}_\zeta\Big[\,w'_a(\mathcal{A}^{D,u},\delta,1) \geq \eta; \mathcal{A}^{D,u}_{\tau_1^D} = \mathcal{A}^{D,u}_{V_1^D} \text{ or } \mathcal{A}^{D,u}_{V_1^D} = \mathcal{A}^{D,u}_{\sigma_2^D} \;\Big|\; \frac{1}{3} \leq \tau_1^D < \frac{1}{2},&\\ \sigma_2^D - \tau_1^D < \frac{\delta}{2},\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_n\Big]&\\ +\mathbf{P}_\zeta\Big[w'_a(\mathcal{A}^{D,u},\delta,1) \geq \eta\ ;\ \mathcal{A}^{D,u}_{\tau_1^D} \neq \mathcal{A}^{D,u}_{V_1^D}\ ;\ \mathcal{A}^{D,u}_{V_1^D} \neq \mathcal{A}^{D,u}_{\sigma_2^D}\Big|\; \frac{1}{3} \leq \tau_1^D < \frac{1}{2},&\\ \sigma_2^D - \tau_1^D < \frac{\delta}{2},\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_n\Big].&\end{aligned} \tag{4.30}$$

The first term in (4.30) is nonnegative, and if we are in the conditions given by the second term, then

$$d\big(\mathcal{A}^{D,u}_{\tau_1^D}, \mathcal{A}^{D,u}_{V_1^D}\big) \geq \epsilon, \quad d\big(\mathcal{A}^{D,u}_{V_1^D}, \mathcal{A}^{D,u}_{\sigma_2^D}\big) \geq \epsilon \quad \text{and} \quad \sigma_2^D - \tau_1^D < \frac{\delta}{2},$$

implying that $w'_a(\mathcal{A}^D,\delta,1) \geq \epsilon > \eta$. Therefore, the second term in (4.30) is equal to

$$\begin{aligned}&\mathbf{P}_\zeta\Big[\mathcal{A}^{D,u}_{\tau_1^D} \neq \mathcal{A}^{D,u}_{V_1^D}\ ;\ \mathcal{A}^{D,u}_{V_1^D} \neq \mathcal{A}^{D,u}_{\sigma_2^D}\Big|\; \frac{1}{3} \leq \tau_1^D < \frac{1}{2},\ \sigma_2^D - \tau_1^D < \frac{\delta}{2},\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_n\Big]\\ &\qquad = \mathbf{P}_\zeta\Big[\mathcal{A}^{D,u}_{\tau_1^D} \neq \mathcal{A}^{D,u}_{V_1^D}\ ;\ \mathcal{A}^{D,u}_{V_1^D} \neq \mathcal{A}^{D,u}_{\sigma_2^D}\Big|\; \frac{1}{3} \leq \tau_1^D < \frac{1}{2},\ \mathcal{A}^D_{\tau_1^D} \notin \Pi_n\Big](1+o(1)).\end{aligned}$$

Now, by the strong Markov property applied to $\mathcal{A}^D$ at time τ_1^D, we have

$$\begin{aligned}(4.31)\qquad &\mathbf{P}_\zeta\Big[\,\mathcal{A}^{D,u}_{\tau_1^D} \neq \mathcal{A}^{D,u}_{V_1^D}\,; \mathcal{A}^{D,u}_{V_1^D} \neq \mathcal{A}^{D,u}_{\sigma_2^D}\,;\, \frac{1}{3} \leq \tau_1^D < \frac{1}{2};\, \mathcal{A}^D_{\tau_1^D} \notin \Pi_n\Big]\\ &= \mathbf{E}_\zeta\Big[\mathbf{P}_{\mathcal{A}^D_{\tau_1^D}}\Big[\mathcal{A}^{D,u}_{\tau_1^D} \neq \tilde{\mathcal{A}}^{D,u}_{\tilde{U}_1^D}\,;\, \tilde{\mathcal{A}}^{D,u}_{\tilde{U}_1^D} \neq \tilde{\mathcal{A}}^{D,u}_{\tilde{\sigma}_1^D}\Big]\,\mathbb{I}_{\{1/3\leq\tau_1^D<1/2\}}\mathbb{I}_{\{\mathcal{A}^D_{\tau_1^D}\notin\Pi_n\}}\Big].\end{aligned}$$

Since we assumed that Condition (i) did not hold, there exists $\eta \in \chi(\Pi_n)\setminus\Pi_n$ such that $\mathbf{P}_\eta[\tilde{U}_1^D < \tilde{\sigma}_1^D] \geq C_1$ for a constant $C_1 > 0$ and D large enough. As $\eta \in \chi(\Pi_n)$, we can choose ζ such that $\mathbb{P}[\chi(\underline{\zeta}) = \eta] > 0$ (and $\mathbb{P}_{\underline{\zeta}}[\sigma_2 < \infty] > 0$). Now, since we assumed that $\mathbf{P}[\underline{\gamma}^u \neq \gamma^u] > 0$ for all $\gamma \notin \Pi_n$, the probability that a coalescence event occurs before a scattering event in the structured genealogical process ξ started at any value not in Π_n is greater than a constant C_2. Therefore, we can write

$$\mathbf{P}_\eta\Big[\eta^u \neq \tilde{\mathcal{A}}^{D,u}_{\tilde{U}_1^D}\,;\, \tilde{\mathcal{A}}^{D,u}_{\tilde{U}_1^D} \neq \tilde{\mathcal{A}}^{D,u}_{\tilde{\sigma}_1^D}\Big] > C_1'$$

for a constant $C_1' > 0$. By the distribution of the epochs of the geographical collisions, the convergence in law of $(\tau_1^D, \mathcal{A}^D_{\tau_1^D})$ to $(\sigma_2, \chi(\underline{\zeta}))$ (c.f. Lemma 4.20) and the fact that $\eta \notin \Pi_n$, we have for ζ chosen as above

$$\mathbf{P}_\zeta\Big[\mathcal{A}^D_{\tau_1^D} = \eta\,;\, \frac{1}{3} \leq \tau_1^D < \frac{1}{2};\, \mathcal{A}^D_{\tau_1^D} \notin \Pi_n\Big] > C_3$$

for a constant $C_3 > 0$ and D large enough, so the expression in the right-hand side of (4.31) is bounded below by $C_1'C_3$, and so is (4.29). Hence, $(ii) \Rightarrow (i)$.

Suppose now that Condition (i) is fulfilled. Condition (a) of Corollary 3.7.4 in [**EK86**] trivially holds, so we only need to check Condition (b) on the modulus of continuity. Fix $\zeta \in \mathrm{P}_n^s$ and $a > 0$, and let $T > a$ and $\eta > 0$. Firstly, by the convergence in probability of σ_1^D to 0, there exists $D_1 \in \mathbb{N}$ such that for all $D \geq D_1$, $\mathbf{P}_\zeta[\sigma_1^D \geq a] < \eta/5$. Secondly, we have

$$\mathbf{P}_\zeta\big[\text{at least one } R_D\text{−event in } [0,T]\big] \leq 1 - \exp\big(-T\max_{\xi\in\mathrm{P}_n^s} c_{R_D}(\xi)\big) \to 0, \qquad D \to \infty,$$

so there exists $D_2 \geq 1$ such that for all $D \geq D_2$, the previous quantity is less than $\eta/5$. Thirdly, by the same argument as in the beginning of the proof of Theorem 4.4, there exists $N \in \mathbb{N}$ such that $\mathbb{P}_{\underline{\zeta}}[\sigma_N \leq T] < \eta/5$. Hence, by Lemma 4.20, there exists $D_3 \geq 1$ such that for all $D \geq D_3$, $\mathbf{P}_\zeta[\sigma_N^D \leq T] < \eta/5$.

Consequently, we can write for each $D \geq \max\{D_1, D_2, D_3\}$ and all $\delta > 0$

$$\begin{aligned}&\mathbf{P}_\zeta\big[\,w_a'\,(\mathcal{A}^{D,u},\delta,T) \geq \eta\big]\\ &\quad\leq \mathbf{P}_\zeta[\sigma_1^D \geq a] + \mathbf{P}_\zeta[\sigma_N^D \leq T] + \mathbf{P}_\zeta\big[\text{at least one } R_D\text{−event in } [0,T]\big]\\ &\qquad\quad + \mathbf{P}_\zeta\big[w_a'(\mathcal{A}^{D,u},\delta,T) \geq \eta;\ \sigma_1^D < a;\ \sigma_N^D > T;\ \text{no } R_D\text{−events in } [0,T]\big]\\ &\quad< \frac{3\eta}{5} + \mathbf{P}_\zeta\big[w_a'(\mathcal{A}^{D,u},\delta,T) \geq \eta;\ \sigma_1^D < a;\ \sigma_N^D > T;\ \text{no } R_D\text{−events in } [0,T]\big].\end{aligned}$$

Furthermore, there exists $\delta > 0$ such that $\mathbb{P}_\gamma[\sigma_2 < 3\delta] < \eta/(5N)$ for all $\gamma \in \Pi_n$. Now, for all $i \in \{1,\dots,N\}$, by the strong Markov property applied to $\mathcal{A}^D$ at time τ_{i-1}^D and the convergence of $\mathbf{P}_\gamma[\tau_1^D < 3\delta]$ to $\mathbb{P}_{\underline{\gamma}}[\sigma_2 < 3\delta]$, uniformly in γ, we have

$$\mathbf{P}_\zeta\big[\tau_{i-1}^D < \infty;\ \tau_i^D - \tau_{i-1}^D < 3\delta\big] = \mathbf{E}_\zeta\big[\mathbb{I}_{\{\tau_{i-1}^D<\infty\}}\,\mathbf{P}_{\mathcal{A}^D_{\tau_{i-1}^D}}\,[\tilde{\tau}_1^D < 3\delta]\big] \leq \frac{\eta}{5N}$$

for D large enough. Therefore,

$$\mathbf{P}_\zeta\big[\,w_a'\,(\mathcal{A}^{D,u},\delta,T) \geq \eta;\ \sigma_1^D < a;\ \sigma_N^D > T;\ \text{no } R_D\text{−event in } [0,T]\big]$$

$$\begin{aligned}&\leq \sum_{i=1}^{N}\mathbf{P}_{\zeta}\big[\tau_{i-1}^{D}<\infty;\ \tau_{i}^{D}-\tau_{i-1}^{D}<3\delta\big]\\ &\quad+\mathbf{P}_{\zeta}\big[w_a'(\mathcal{A}^{D,u},\delta,T)\geq\eta;\ \sigma_1^D<a;\ \sigma_N^D>T;\ \text{no } R_D-\text{event in } [0,T];\\ &\qquad\qquad \tau_i^D-\tau_{i-1}^D\geq 3\delta \text{ for all } i\leq N \text{ s.t. } \tau_{i-1}^D<\infty\big]\end{aligned}\tag{4.32}$$

and the first sum is less than $\eta/5$. To finish, let V_i^D denote the epoch of the next event after τ_i^D if $\mathcal{A}^D_{\tau_i^D}\notin\Pi_n$ (if it exists, $V_i^D=+\infty$ otherwise), and set $V_i^D=\tau_i^D=\sigma_{i+1}^D$ if $\mathcal{A}^D_{\tau_i^D}\in\Pi_n$. Since we assume that Condition (i) holds, for all $i\in\{1,\ldots,N\}$ we have by the strong Markov property applied at time τ_i^D and the fact that the distribution of $\mathcal{A}^D_{\tau_i^D}$ concentrates on $\chi(\Pi_n)$ as D grows to infinity by Lemma 4.20,

$$\mathbf{P}_{\zeta}\big[\tau_i^D<\infty;V_i^D<\sigma_{i+1}^D\big]\to 0,\qquad D\to\infty,$$

so the last term in (4.32) is less than

$$\begin{aligned}&\sum_{i=1}^{N}\mathbf{P}_{\zeta}\left[\tau_i^D<\infty;V_i^D<\sigma_{i+1}^D\right]\\ &\quad+\ \mathbf{P}_{\zeta}\big[w_a'(\mathcal{A}^{D,u},\delta,T)\geq\eta;\ \sigma_1^D<a;\ \sigma_N^D>T;\ \text{no } R_D-\text{event in } [0,T];\ \tau_i^D-\tau_{i-1}^D\geq 3\delta\\ &\qquad\qquad \text{and } V_i^D=\sigma_{i+1}^D \text{ for all } i\leq N \text{ s.t. } \tau_{i-1}^D<\infty\big],\end{aligned}$$

where the first sum is less than $\frac{\eta}{5}$ for D large enough. But on that last event, σ_1^D is less than a and no R_D-events occur so τ_1^D is the epoch of the event directly after σ_1^D, then all geographical collisions are at least 3δ away from each other and the σ_i^D's are the only times in between at which an event occurs, so necessarily $w_a'(\mathcal{A}^D,\delta,T)=0$. Assembling all the pieces, we obtain that

$$\mathbf{P}_{\zeta}\big[w_a'(\mathcal{A}^{D,u},\delta,T)\geq\eta\big]<\eta,$$

completing the proof of $(i)\Rightarrow(ii)$.

If $\zeta\in\chi(\Pi_n)\cup\Pi_n$, then we only need to show that (i) implies the tightness of $(\mathcal{A}^{D,u})_{D\geq 1}$ on $[0,\infty)$. Let us directly use the same notation as in the last proof. In the last paragraph, we proved that with a high probability, there is no accumulations of jumps between the random time τ_1^D and T. Also, we can make $\mathbf{P}_{\zeta}[\tau_1^D\leq 2a]$ as small as we want by adjusting a and taking D large enough, and the probability that at least one R_D-event occurs is vanishingly small, so we are left with proving that, if δ is such that $\mathbf{P}_{\zeta}[w_a'(\mathcal{A}^{D,u},\delta,T)\geq\eta]<\eta$, $\tau_1^D>2a$ and no R_D-events occur between 0 and T, then $\mathbf{P}_{\zeta}[w'(\mathcal{A}^{D,u},\delta',a)\geq\eta]<\eta$, for some $\delta'\in(0,\delta)$. If $\zeta\in\Pi_n$, $\tau_1^D>2a$ and no R_D-events occur, then τ_1^D is the epoch of the first event occurring to $\mathcal{A}^D$ so $w'(\mathcal{A}^{D,u},\delta',a)=0$ for all $\delta'\in(0,\delta\wedge a)$. If $\zeta\in\chi(\Pi_n)$ and the other conditions hold, then by Condition (i) we have

$$\mathbf{P}_{\zeta}\big[U_1^D=\sigma_1^D\big]\to 1,\qquad D\to\infty,$$

and furthermore $\mathbf{P}_{\zeta}[\sigma_1^D<a]\to 1$, so with a probability tending to one as D grows to infinity, one event occurs between 0 and a, then nothing happens between a and $2a$ (there is no R_D-events, so the next event after σ_1^D must occur at time $\tau_1^D>2a$) and the condition on the modulus of continuity is fulfilled after time a so, for any $\delta'\in(0,\delta\wedge a)$, we do have

$$\mathbf{P}_{\zeta}\big[w'(\mathcal{A}^{D,u},\delta',a)\geq\eta\big]<\eta.$$

This completes the proof of the case $\zeta\in\chi(\Pi_n)\cup\Pi_n$.

Now, by Theorem 3.7.8 in [**EK86**], the two ingredients to obtain the convergence of the processes $(\mathcal{A}^D)_{D\geq 1}$ are tightness, given by the first part of Proposition 4.23 for any $a>0$, and convergence of the finite-dimensional distributions, given by Theorem 4.4 and the bijective correspondence between Π_n and $\mathcal{P}_n$. For $\zeta \in \Pi_n$ and $a = 0$, tightness still holds by virtue of the last paragraph, and an easy modification (namely allowing $t = 0$ in the proof of the convergence of the one-dimensional distributions) of the proof of Theorem 4.4 in that case, where $\underline{\zeta} = \zeta$ and $\sigma_1^D = 0$ a.s., gives the convergence of the finite-dimensional distributions of $\mathcal{A}^D$, including at time $t = 0$. $\square$

Let us briefly comment on the condition $\mathbf{P}_\xi[U_1^D = \sigma_1^D] \to 1$. If the fast within-deme coalescence is given by a Ξ-coalescent (including Kingman's coalescent) occurring in one deme at a time, the condition is fulfilled if and only if at most two lineages can be collected into the same deme during a single event. Indeed, in that case the next step of the genealogical process is either to scatter these two lineages into two different demes or to merge them into one lineage, the outcome of which is always in Π_n. If more than 2 lineages are gathered in the same deme and do not merge during the geographical collision, then with a positive probability only two of them are involved in the next genealogical event and at least two rapid steps are needed for $\mathcal{A}^D$ to re-enter Π_n. The same conclusion holds if two pairs of lineages are gathered in two demes (meaning 2 lineages per deme), since the genealogical process acts in one deme at a time by assumption.

4. Collapse of structured genealogical processes

The next proposition states that the only reasonable structured genealogies which collapse to an unstructured genealogy (given by a Ξ-coalescent) when the number of demes tends to infinity are the genealogies that we have described before, subject to certain conditions.

Note that if we want the lineages to be exchangeable in the limit, the limiting process needs to take its values in $\bigcup_{n\geq 1}\Pi_n$. Indeed, since the rates of intra- and inter-demes mergers greatly differ, we should observe only inter-demes events on the slow time scale. This requires that each deme contains at most one lineage at any given time in the limit.

PROPOSITION 4.25. *Let* $(\mathcal{A}_t^D,\ t \geq 0)_{D\geq 1}$ *be a sequence of structured genealogical processes with values in* $\bigcup_{n\geq 1}\mathrm{P}_n^s$. *Then the following are equivalent*

1. *There exists a sequence* r_D *such that* $r_D \to +\infty$ *as* $D \to \infty$ *and two structured genealogical processes,* $(\xi_t, t \geq 0)$ *(resp.* $(\mathcal{A}_t, t \geq 0)$*) with values in* $\bigcup_{n\geq 1}\mathrm{P}_n^s$ *(resp.* $\bigcup_{n\geq 1}\Pi_n$*) satisfying*
   - (a) *for each* $n \in \mathbb{N}$*, the sequence of structured genealogical processes* $(\mathcal{A}^D_{r_D^{-1}\cdot})_{D\geq 1}$ *on the fast time scale, with initial value in* P_n^s*, converges to* ξ *as a process in* $D_{\mathrm{P}_n^s}[0,\infty)$*. In addition,* ξ *is consistent in the sense of Lemma 4.12;*
   - (b) *the sequence* $(\mathcal{A}_t^D, t \geq 1)_{D\geq 1}$ *on the slow time scale converges towards* $\mathcal{A}$ *in that the finite-dimensional distributions (except possibly at time* 0*) converge as in Theorem 4.4 for every sample size* n*;*
   - (c) *there exists a* Ξ*-coalescent* $(R_t, t \geq 1)$ *such that for all* $n \geq 1$*, the unstructured genealogical process* $\mathcal{A}^u$ *induced by* $\mathcal{A}|_{\Pi_n}$ *has the law of the restriction of* R *to* $\mathcal{P}_n$*.*
2. *The rates associated to* $\mathcal{A}^D$ *satisfy Conditions (1), (2) and (3) of Section 3.1, and Condition* (i) *of Lemma 4.12 holds.*

We shall see in the proof that the consistency of ξ is a key ingredient to obtain the desired equivalence. In fact, if we did not impose it, it would certainly be possible to construct particular examples in which the unstructured genealogy on the slow time scale is also a Ξ-coalescent, but the genealogies within a deme are not consistent. We would need to impose 'good' values for the corresponding rates. We rather chose here to emphasize more biologically relevant models, for which the within-deme genealogical process is also consistent and which can be described as part of an entire class of models rather than special cases.

Proof. The implication $2 \Rightarrow 1$ in a consequence of Theorem 4.4, Proposition 4.18 and Proposition 4.16. Let us prove that $1 \Rightarrow 2$. From the definition of a structured genealogical process, blocks can only move and coalesce. Furthermore $\mathcal{A}^D$ stays in P_n^s whenever its initial value lies in this set, so we just need to fix $n \geq 0$ and look at the corresponding rates of scattering, gathering and coalescence. From the description of the limiting processes ξ and $\mathcal{A}$, the rates of $\mathcal{A}^D$ must be of the form

$$r_D \rho_D^{(1)}(\eta|\zeta) + \rho_D^{(2)}(\eta|\zeta) + o(1),$$

where for $i \in \{1, 2\}$, $\rho_D^{(i)}(\eta|\zeta) \to \rho^{(i)}(\eta|\zeta)$ as D tends to infinity. (To simplify notation, we shall write $\rho_D^{(i)}(\eta|\zeta) := \rho^{(i)}(\eta|\zeta)$.) Thus, $\rho^{(1)}(\eta|\zeta)$ are the rates associated to the generator Ψ of the process ξ. Let us check that all cited conditions necessarily hold:

- If $\zeta \to \eta$ is a 1-event, then by adding an $(n+1)$-st individual in one of the existing blocks (therefore changing the sizes of the blocks but not their number), we see that the consistency of ξ imposes that the part of the rate corresponding to the fast time scale depends neither on n, nor on the sizes of the blocks. By exchangeability of the demes, this rate is thus characterized by the number of lineages present in each deme before and after the transition, the order of these numbers being irrelevant. Therefore, Condition (1) holds. By Lemma 4.12, the consistency of ξ implies that Condition (i) of Lemma 4.12 is also satisfied.
- Once again by consistency of ξ, the rate of a 2-event must be of order 1. Indeed, it may otherwise lead to an additional 1-event for the restriction of the process with the $(n+1)$-st lineage (if this additional lineage lands in a non-empty deme or in the same deme as another moving lineage coming from a different subpopulation , and the other dispersing lineages land in different demes), or involve at least two lineages alone in their demes on the fast time scale. If such an event was allowed, then by exchangeability of the islands the fast dynamic could act on a structured partition in Π_n and merge two lineages starting from different demes. Again by exchangeability, any pair of lineages could merge on the fast time scale and so the outcome of ξ would be a single lineage with probability one, a trivial situation which is of no interest here. Now, since we want to keep exchangeability of the lineages in the unstructured genealogy (on the slow time scale), the rates of 2-events should depend only on the number of lineages and their geographical distribution (and possibly on n). But if $\zeta \in \Pi_n$, all lineages are in different demes, so the corresponding rates are necessarily of the form given in Condition (2). If the rates were to depend on n, then as the rates of the fast genealogical process which follows directly (for D large enough, as in the proof of Theorem 4.4) are independent of n, the overall transition from $\eta \in \Pi_n$ to the value of $\mathcal{A}^D$ when it reenters Π_n would eventually give different rates for $\mathcal{A}$ acting on Π_n and for the restriction to Π_n of $\mathcal{A}$ acting on Π_{n+1} (recall the convergence of τ_{i-1}^D and σ_i^D towards σ_i to see that the transitions of $\mathcal{A}$ actually can be described as in Section 3.2). This would contradict the fact that the process $\mathcal{A}^u$ corresponds to a Ξ-coalescent. Finally, we obtain that Condition (2) must hold.
- The last argument imposes also that geographical collisions involving k lineages occur at a rate which is the sum of all corresponding geographical events involving $k+1$ lineages, which is exactly writing the consistency equations (4.4) of Condition (3).

Finally, we obtain that $2 \Rightarrow 1$. □

5. Example

We now turn our attention to a particular class of metapopulation models which combine a (finite) Λ-coalescent within demes with migration between demes and sporadic mass extinction events. We will use the results derived in the preceding sections to characterize the form that the genealogy takes in the infinitely many demes limit. This, in turn, will allow us to illustrate how the statistics of the population-wide Ξ-coalescent depend on the interplay between extinction/recolonization events and the local demographic processes occurring within demes. While these models are quite contrived - in particular, we have simply imposed the condition that a small number of demes is responsible for repopulating vacant demes following a mass extinction - they will allow us to explicitly calculate some quantities of interest.

We first describe how the population evolves forwards-in-time. Suppose that for each D, each deme contains exactly N individuals. Fix $K \in \mathbb{N}$, and let $\Lambda^d(dx)$ and $\Lambda^g(dy)$ be two probability measures on $[0,1]$ with no atom at 0. Then reproduction, migration, and extinction/recolonization events occur according to the following rules.

- Each individual in each deme reproduces at rate D according to the following scheme. If an individual in deme i reproduces, then a number x is sampled from $[0,1]$ according to the probability distribution $\Lambda^d(dx)$, and then each occupant of that deme dies with probability x and is replaced by an offspring of the reproducing individual. In terms of the notation of Section 2.1, such an event has the following representation when k is the label of the reproducing individual. First, $R^{j,j'} = (0,\dots,0)$ for all pairs of integers $j \neq j' \in [D]$ and $R^{j,j} = (1,\dots,1)$ if $j \in [D] \setminus \{i\}$. $R^{i,i}$ is a random vector obtained by choosing a number x according to $\Lambda^d(dx)$, a number m according to a binomial distribution with parameters (N,x), and finally a set $\mathcal{O} \subset [N]$ of offspring of the reproducing individual by sampling m labels in $[N]$ uniformly without replacement. Then, $R^{i,i}_k = m$, $R^{i,i}_{k'} = 0$ for all $k' \in \mathcal{O} \setminus \{k\}$, and $R^{i,i}_l = 1$ for all $l \notin \mathcal{O} \cup \{k\}$.
- At rate Dm_1, each individual gives birth to a single migrant offspring which then moves to any one of the D demes, chosen uniformly at random, and replaces one of the N individuals within that deme, also uniformly at random. In this case, if j is the label of the deme containing the parent and k is its label, then a pair (i,l) is sampled uniformly in $[D] \times [N]$ and the vectors R are as described in Example 4.7 of Section 2.1.
- Mass extinction events occur at rate e. When such an event occurs, a number y is sampled from $[0,1]$ according to the probability distribution $\Lambda^g(dy)$. Then, each deme goes extinct with probability y, independently of all the others, and is unaffected by the extinction otherwise. Simultaneously, K of the D demes are chosen uniformly at random to be source demes, and the deceased occupants of the extinct island are replaced by offspring produced by individuals living in the source demes according to the following scheme. The parent of each individual recolonizing a deme left vacant by the mass extinction is chosen uniformly at random and with replacement from among the NK inhabitants of the source demes. If a source deme is chosen from among the extinct ones, then the parents of the offspring emerging from that deme are the individuals that occupied the deme immediately prior to the extinction. To describe such an event using the notation of Section 2.1, suppose that a number y is chosen according to $\Lambda^g(dy)$, a number m is sampled according to a $\mathrm{Binom}(D,y)$-distribution and a (random) set $\mathcal{O}_{\text{ext}} \subset [D]$ is constructed by sampling uniformly without replacement m deme labels. Independently, another set $\mathcal{O}_{\text{rec}}$ of K recolonizing demes is also chosen by uniform sampling. Then, for all $i \notin \mathcal{O}_{\text{ext}}$ we have $R^{i,i} = (1,\dots,1)$ and each deme $j \in \mathcal{O}_{\text{ext}} \setminus \mathcal{O}_{\text{rec}}$ satisfies $R^{i,j} = (0,\dots,0)$ for all $i \in [D]$.

The vectors $R^{i,j}$ with $j \in \mathcal{O}_{\mathrm{rec}}$ and $i \in \mathcal{O}_{\mathrm{ext}} \cup \mathcal{O}_{\mathrm{rec}}$ are not easily formulated explicitly (in particular, their description depends on whether the recolonizing demes also go extinct during the event), but it is clear that the evolution of the population satisfies the two conditions required in Section 2.1.

Suppose that n individuals are sampled from the population at time 0, and let us consider the evolution (backwards-in-time) of the structured coalescent process $\mathcal{A}^D$ in P_n^s. From the description of the model forwards-in-time, the events affecting the genealogy occur at the following rates:

1. If a deme contains b lineages, then each k-tuple of lineages in this deme (for $k \leq b$) merges into one lineage in the same deme at rate
$$D\lambda^d_{b;k,1,\dots,1} = DN \int_0^1 \Lambda^d(dx) x^k (1-x)^{b-k}.$$
Furthermore, any merger event occurs in one deme at a time.
2. Each lineage migrates (alone) at rate Dm_1. Indeed, the total rate at which migrant offspring are produced forwards-in-time is $ND \times Dm_1$, but the probability that such a migrant belongs to the lineage under consideration is $(ND)^{-1}$ (recall that the deme and the label of the individual replaced by the migrant are chosen uniformly at random). Consequently, the probability that a migrating lineage lands in a non-empty deme is D^{-1} times the number of demes occupied by the other lineages of $\mathcal{A}^D_{t-}$. When such an event occurs, the probability that the migrating lineage also merges with an ancestral lineage present in the source deme is N^{-1} times the number of distinct ancestral lineages present in that deme.
3. Extinction events generate geographical collisions at rate $\mathcal{O}(1)$. Because the K recolonizing demes are chosen uniformly from among the D islands, recolonization by a deme containing at least one lineage of the genealogical process occurs with a probability of order $\mathcal{O}(D^{-1})$, and so these events are negligible in the limit. Suppose that $\mathcal{A}^D_{t-} \in \Pi_n$. Let $k \leq |\mathcal{A}^D_{t-}|$, $r \leq K$, and let $k_1, \dots, k_r$ be integers greater than 1 and summing to k. For each $i \in \{1, \dots, r\}$, let $L_i = \{l_{i1}, \dots, l_{ij_i}\}$ be a collection of j_i integers summing to k_i. Then each $(|\mathcal{A}^D_{t-}|; k_1, \dots, k_r, 1, \dots, 1;\, L_1, \dots, L_r, \{1\}, \dots, \{1\})$-geographical collision occurs at rate

$$\begin{aligned} & e \int_0^1 \Lambda^g(dy) \sum_{s=0}^{|\mathcal{A}^D_{t-}|-k} \mathbb{I}_{\{s \leq K-r\}} \binom{|\mathcal{A}^D_{t-}|-k}{s} y^{k+s}(1-y)^{|\mathcal{A}^D_{t-}|-k-s} \frac{K!}{(K-r-s)!} \frac{1}{K^{k+s}} \\ & \qquad \times \prod_{i=1}^{r} \left\{ \frac{N!}{(N-j_i)!} \frac{1}{N^{k_i}} \right\} + \mathcal{O}\Big(\frac{1}{D}\Big) \qquad (4.33) \\ & =: \tilde{\lambda}^g_{|\mathcal{A}^D_{t-}|;k_1,\dots,k_r,1,\dots,1;L_1,\dots,L_r,\{1\},\dots,\{1\}} + \mathcal{O}\Big(\frac{1}{D}\Big). \end{aligned}$$

The rate expression that appears in (4.33) can be interpreted in the following way. As well as the k ancestral lineages that are known to be affected by the disturbance (this is specified by the type of event), an additional s lineages may be caught up in the extinction event and moved to demes where they remain isolated (hence producing no changes in the structured genealogy). In (4.33), the first part in each term of the sum corresponds to the number of choices for these s additional lineages, followed by the probability that only these $k+s$ lineages are affected. The condition $r+s \leq K$ is imposed by the fact that the r groups of lineages geographically gathered and the s lineages affected but remaining alone in their demes must then belong to $r+s$ distinct recolonizing demes. The middle part of the term specifies the probability that the affected lineages are grouped in the desired way: regardless of the labels of the recolonizing demes, if the latter contain no lineages of the sample just before the extinction then $\frac{K!}{(K-r-s)!}$ is the number of (unordered) ways of choosing $r+s$ of them to receive the affected lineages, while K^{-k-s} is the probability that each of the $k+s$ lineages moves to the prescribed recolonizing deme. Finally, the last product is obtained in a similar manner by allocating as many distinct ancestors as required to the groups of lineages gathered into the same demes. As explained above, the $\mathcal{O}(D^{-1})$ remainder term accounts for the probability that at least one of the finitely-many recolonizing demes contains a lineage of $\mathcal{A}^D_{t-}$.

Let us say that a *simple collision* occurs when a single lineage moves into a non-empty deme, and possibly merges with one of the lineages present in this deme. To verify that the convergence results from the previous sections apply to the example, it will be convenient to introduce the following quantities, defined for all $\zeta, \eta \in \mathrm{P}^s_n$:

$$\phi_c(\zeta,\eta) = \begin{cases} 1 & \text{if } \zeta \to \eta \text{ is a simple collision with coalescence,} \\ 0 & \text{otherwise,} \end{cases}$$

and likewise

$$\phi_{nc}(\zeta,\eta) = \begin{cases} 1 & \text{if } \zeta \to \eta \text{ is a simple collision without coalescence,} \\ 0 & \text{otherwise.} \end{cases}$$

By 'with coalescence' (resp. 'without coalescence'), we mean that the migrating lineage merges (resp. does not merge) with a lineage in the source deme during the same event.

Let us consider a particular 1-event. If this event involves a single lineage moving to an empty deme, it may be caused either by a migration event of the kind described in item 2 above (which occurs at rate $Dm_1(1-k/D)$ if k is the number of demes occupied by the other lineages at the time of the event), or by a mass extinction event (whose rate is of order $\mathcal{O}(1)$ according to item 3). Consequently, the overall rate of any 1-event is of the form $Dm_1 + \mathcal{O}(1)$. Groups of more than one lineage can also move simultaneously, but only through an extinction event and so at a rate of order $\mathcal{O}(1)$. If the event involves an intra-deme merger, then its rate is easily written in the form given in Section 3.1 with $r_D = D$; see item 1. A 2-event $\zeta \to \eta$ occurs at a rate of order $\mathcal{O}(1)$, and in particular if $\zeta \in \Pi_n$, then this rate is given by

$$\tilde{\lambda}^g(\zeta,\eta) + 2m_1\Big\{\phi_c(\zeta,\eta)\frac{1}{N} + \phi_{nc}(\zeta,\eta)\frac{N-1}{N}\Big\} + \mathcal{O}\Big(\frac{1}{D}\Big) =: \lambda^g(\zeta,\eta) + \mathcal{O}\Big(\frac{1}{D}\Big),$$

where $\tilde{\lambda}^g(\zeta,\eta)$ is the rate of the unique extinction event which turns ζ into η. In this expression, the term in brackets is nonzero only if the event is a simple collision involving two lineages that have been collected in the same deme through migration. Such collisions occur at rate $2m_1$, and then the two lineages either coalesce, with probability N^{-1}, or remain distinct, with probability $1-N^{-1}$. Finally, we must check that the λ^g's satisfy (4.4), and that the rates on the fast time scale satisfy Condition (i) of Lemma 4.12. The latter condition follows from the description of the rates and the consistency of Λ-coalescents, and the validity of the former condition can be deduced from the fact that lineages choose independently of each other whether they are involved in the event, and which of the NK individual they take as a parent. We leave the straightforward but tedious calculations to the interested reader.

All conditions of Theorem 4.4 and Proposition 4.16 hold. Thus, we can conclude that the finite dimensional distributions of $\mathcal{A}^D$ converge to those of a structured genealogical process $\mathcal{A}$ with values in Π_n, and that the unstructured process $\mathcal{A}^u$ is a Ξ-coalescent with values in $\mathcal{P}_n$. Let us describe $\mathcal{A}^u$ as precisely as we can. To apply the results of Section 2.3, we need to know the distribution of the final state of the 'fast' process ξ that was introduced in Section 2.2. Starting from a structured partition where all blocks are contained in the same component (i.e., all lineages lie initially in the same deme), this distribution coincides with the sampling distribution of the infinitely many alleles model of the generalized Fleming-Viot process dual to the Λ-coalescent with finite measure $x^2\Lambda^d(dx)$ acting within this deme. Indeed, on the fast time scale, ancestral lineages belonging to a common deme migrate out to distinct, empty islands, a process analogous to mutation to unique types with a 'mutation' rate equal to m_1. Recursion formulae are given in [**Möh06**] which can be used to compute the probability $p(\mathbf{n})$ of unordered allele configurations $\mathbf{n} = \{n_1, \dots, n_k\}$ in the infinitely many alleles model when the genealogy is given by a Λ- or a Ξ-coalescent. In our case, the formula of interest is (with $p(1) = 1$):

$$p(\mathbf{n}) = \frac{nm_1}{g_n + nm_1} \sum_{j=1}^{k} \mathbb{I}_{\{n_j=1\}} \frac{1}{k}\, p(\tilde{\mathbf{n}}_j) + \sum_{i=1}^{n-1} \frac{g_{n,n-i}}{g_n + nm_1} \sum_{j=1}^{k} \mathbb{I}_{\{n_j>i\}} \frac{n_j - i}{n-i}\, p(\mathbf{n} - i\mathbf{e}_j),$$

where $n := \sum_{j=1}^{k} n_j \geq 2$, $\tilde{\mathbf{n}}_j := (n_1, \dots, n_{j-1}, n_{j+1}, \dots, n_k)$, $\mathbf{e}_j$ denotes the j'th unit vector in $\mathbf{R}^k$ and g_{nk} (resp. g_n) is the rate at which the number of lineages decreases from n to k (resp. the total rate at which the number of lineages changes when n lineages are alive), given by

$$g_{nk} = \binom{n}{k-1} \int_0^1 \Lambda^d(dx) x^{n-k+1}(1-x)^{k-1}$$

and

$$g_n = \sum_{k=1}^{n-1} g_{nk} = \int_0^1 \Lambda^d(dx)\big(1 - (1-x)^{n-1}(1 - x + nx)\big).$$

These expressions are related to the distribution of $\underline{\zeta}$ by the following formula:

$$\mathbf{P}\big[\underline{\zeta} = \big(\{B_1\}, \dots, \{B_k\}, \emptyset, \dots, \emptyset\big)\big] = p(|B_1|, \dots, |B_k|),$$

where $\zeta = \big(\{\{1\}, \dots, \{n\}\}, \emptyset, \dots, \emptyset\big)$ and $|B_i|$ denotes the number of elements in the block B_i. Indeed, because the dynamics on the fast time scale of lineages occupying different demes are independent, the final state of the fast genealogical process is the concatenation of all the final states of the groups of lineages starting in the same deme. Hence, the preceding results are sufficient to describe $\underline{\zeta}$ for any $\zeta \in \mathrm{P}_n^s$. Unfortunately, with this level of generality, there does not appear to be a simple description of the measure Ξ associated to $\mathcal{A}^u$, but the rate associated to its Kingman part (that is its mass at $\mathbf{0}$) is given by:

$$2m_1\frac{1}{N} + 2m_1\frac{N-1}{N}\, p(2) = 2\frac{m_1}{N}\left\{1 + (N-1)\, \frac{\int_0^1 \Lambda^d(dx)x^2}{\int_0^1 \Lambda^d(dx)x^2 + 2m_1}\right\}. \tag{4.34}$$

The first term in (4.34) corresponds to a simple collision with coalescence, and the second term to a simple collision without coalescence; the probability that the lineages then coalesce before one of them migrates is given by $p(2)$.

One case which can be characterized more thoroughly is when dispersal between demes only occurs during extinction-recolonization events ($m_1 = 0$). For example, this might be a reasonable approximation to make when modeling a population in which migrants are at a substantial competitive disadvantage relative to residents, so that dispersal is only successful into demes in which the resident population has gone extinct. In this case, the Kingman component of the genealogy disappears (see (4.34)). Furthermore, viewed backwards in time, lineages gathered into common demes by mass extinction events cannot migrate away before the rapid within-deme coalescent reaches a common ancestor, and so any such group of lineages merges instantaneously into a single lineage. The shape of the resulting global coalescent therefore is determined only by the way in which mass extinction events gather lineages together. Recall the expression for the rates of geographical collisions given in (4.33), and let us examine how K, the number of demes contributing colonists in the wake of a mass extinction, affects the shape of the genealogy.

If $K = 1$, all lineages affected by a mass extinction event have parents within the same deme. The resulting genealogy is a Λ-coalescent, and the rate at which k ancestral lineages merge when m are present is equal to the rate at which exactly k lineages are caught up in an extinction event when m demes contain one lineage, that is

$$e\int_0^1 \Lambda^g(dy)y^k(1-y)^{m-k}.$$

On the other hand, if we let K tend to infinity, then each term in the sum in (4.33) is asymptotically equivalent to $\frac{K!}{(K-r-s)!}\,K^{-k-s} \sim K^{r-k}$, up to a constant (recall that the sample size n is finite and bounds the number of lineages at any times). Consequently, binary geographical collisions ($k = k_1 = 2$, $r = 1$, $j_1 \in \{1,2\}$) occur at a rate of order $\mathcal{O}(K^{-1})$, whereas the rate of a collision involving at least 3 lineages is of order at most $\mathcal{O}(K^{-2})$. Hence, for fixed sample size n, the probability that only binary mergers occur in the sample genealogy approaches 1 as K tends to infinity, and the rate of each binary merger (multiplied by K) converges to

$$e\int_0^1 \Lambda^g(dy)y^2, \tag{4.35}$$

where the term y^2 is obtained by observing that the condition $s \leq K-1$ in (4.33) is always fulfilled for n fixed and K large enough, and that $\sum_{s=0}^{|\mathcal{A}^D_{t-}|-2}\binom{|\mathcal{A}^D_{t-}|-2}{s}y^{2+s}(1-y)^{|\mathcal{A}^D_{t-}|-2-s} = y^2$. Once the lineages are gathered into the same deme, they can only coalesce and they do so instantaneously on the slow time scale as $D \to \infty$. It follows that if time is rescaled by a factor of DK, then the rate of a binary merger converges to that of Kingman's coalescent run at the rate shown in (4.35). Moreover, under this time rescaling, the rates of the finitely many possible multiple merger events converge to 0 as K grows to infinity, and so the limiting (as $D \to \infty$) unstructured genealogical process $\mathcal{A}^u$ corresponding to an evolution with K recolonizing demes converges to Kingman's coalescent as a process in $D_{\mathcal{P}_n}[0,\infty)$ as K tends to infinity. (Note, however, that this does not imply that one can interchange the limits $D \to \infty$ and $K \to \infty$.) Finally, if K is finite but greater than 1, then geographical collisions involving more than two lineages occur at a non-negligible rate, and so the resulting unstructured genealogy is a more general Ξ-coalescent.

This example shows that a large class of coalescent processes can arise in the infinitely many demes limit of a subdivided population with sporadic mass extinctions. Depending on both the migration and the extinction rates, as well as on the number of demes contributing to population recovery following a mass extinction, the limiting genealogical process can range from Kingman's coalescent ($K = \infty$), as derived by [**Wak04**], to a Λ-coalescent ($K = 1$, $m_1 = 0$), with a family of Ξ-coalescents interpolating between these two extremes. In this particular class of models, multiple mergers of ancestral lineages are more likely to occur when all three parameters, K, N and m_1, are small, so that mass extinctions have a non-negligible probability of gathering lineages into a common deme where they undergo a series of rapid mergers before being scattered again by migration. This observation suggests that it is a generic property of structured population models that if the limiting coalescent admits any multiple mergers, then it also admits simultaneous mergers.

CHAPTER 5

A new model for evolution in a spatial continuum

We introduce a new model for populations evolving in a spatial continuum. This model can be thought of as a spatial version of the Λ-Fleming-Viot process. It explicitly incorporates both small scale reproduction events and large scale extinction-recolonisation events. The lineages ancestral to a sample from a population evolving according to this model can be described in terms of a spatial version of the Λ-coalescent. Using a technique of [**Eva97**], we prove existence and uniqueness in law for the model. We then investigate the asymptotic behaviour of the genealogy of a finite number of individuals sampled uniformly at random (or more generally 'far enough apart') from a two-dimensional torus of side L as $L \to \infty$. Under appropriate conditions (and on a suitable timescale) we can obtain as limiting genealogical processes a Kingman coalescent, a more general Λ-coalescent or a system of coalescing Brownian motions (with a non-local coalescence mechanism). These results are a joint work with N.H. Barton from the Institute of Science and Technology in Vienna, and A.M. Etheridge (cf. [**BEV09**]).

Contents

1. Introduction

In 1982, Kingman introduced a process called the *coalescent*. This process provides a simple and elegant description of the genealogical (family) relationships amongst a set of neutral genes in a randomly mating (biologists would say *panmictic*) population of constant size. Since that time, spurred on by the flood of DNA sequence data, considerable effort has been spent extending Kingman's coalescent to incorporate things like varying population size, natural selection and spatial (and genetic) structure of populations. Analytic results for these coalescent models can be very hard to obtain, but it is relatively easy, at least in principle, to simulate them and so they have become fundamental tools in sequence analysis. However, models of spatial structure have largely concentrated on subdivided populations and a satisfactory model for the ancestry of a population evolving in a two-dimensional spatial continuum has remained elusive. Our aim in this chapter is to present a new model that addresses some of the difficulties of existing models for spatially extended populations while retaining some analytic tractability. The rest of this introduction is devoted to placing this research in context. The reader eager to skip straight to the model and a precise statement of our main results should proceed directly to Section 2.

Our concern here is with the extension of the coalescent to spatially structured populations. In this setting it is customary to assume that the population is subdivided into *demes* of (large) constant size, each situated at a vertex of a graph G, and model the genealogical trees using the *structured* coalescent. As we trace backwards in time, within each deme the ancestral lineages follow Kingman's coalescent, that is each pair of lineages merges (or *coalesces*) into a single lineage at a constant rate, but in addition lineages can migrate between demes according to a random walk on the graph G. The genealogical trees obtained in this way coincide with those for a population whose forwards in time dynamics are given by Kimura's stepping stone model (see [**Kim53**]) or, as a special case, if G is a complete graph, by Wright's island model introduced in [**Wri31**].

The stepping stone model is most easily described when the population consists of individuals of just two types, a and A say. It can be extended to incorporate selection, but let us suppose for simplicity that these types are selectively neutral. Labelling the vertices of the graph G by the elements of the (finite or countable) set I and writing p_i for the proportion of individuals in deme i of type a, say, we have

$$dp_i(t) = \sum_{j \in I} m_{ji} \left(p_j(t) - p_i(t)\right) dt + \sqrt{\gamma p_i(t) \left(1 - p_i(t)\right)}\, dW_i(t) \tag{5.1}$$

where $\{W_i(t); t \geq 0\}_{i \in I}$ is a collection of independent Wiener processes, γ is a positive constant and $\{m_{ij}\}_{i,j \in I}$ specifies the rates of a continuous time random walk on G. The graph G, chosen to caricature the spatial structure of the population, is typically taken to be $\mathbb{Z}^2$ (or its intersection with a two-dimensional torus) and then one sets $m_{ij} = \kappa\, \mathbb{I}_{\{\|i-j\|=1\}}$, corresponding to simple random walk.

Although the stepping stone model is widely accepted as a model for structured populations, in reality many populations are not subdivided, but instead are distributed across a spatial continuum. Wright in [**Wri43**] and Malécot in [**Mal48**] derived expressions for the probability of identity of two individuals sampled from a population dispersed in a two dimensional continuum by assuming on the one hand that genes reproduce and disperse independently of one another, and on the other hand that they are scattered in a stationary Poisson distribution. However, these assumptions are incompatible ([**Fels75, SF79**]). The assumption of independent reproduction will result in 'clumping' of the population and some local regulation will be required to control the local population density.

A closely related approach is to assume that the genealogical trees can be constructed from Brownian motions which coalesce at an instantaneous rate given by a function of their separation. The position of the common ancestor is typically taken to be a Gaussian centred on the midpoint between the two lineages immediately before the coalescence event (although other distributions are of course possible). However, the coalescent obtained in this way does not exhibit *sampling consistency*. That is, if we construct the genealogical tree corresponding to a sample of size n and then examine the induced genealogical tree for a randomly chosen subsample of size $k < n$, this will not have the same distribution as the tree we obtain by constructing a system of coalescing lineages directly from the subsample. The reason is that whenever one of the lineages in the subsample is involved in a coalescence event in the full tree it will jump. Furthermore, just as in Malécot's setting, there is no corresponding *forwards* in time model for the evolution of the population.

In [**BDE02**], Barton et al extend the formulae of Wright and Malécot to population models which incorporate local structure. The probability of identity is obtained from a recursion over timeslices of length Δt. Two related assumptions are made. First, the ancestral lineages of genes that are sufficiently well separated are assumed to follow independent Brownian motions (with an effective dispersal rate which will in general differ from the forwards in time dispersal rate) and their chance of coancestry in the previous timeslice is negligible. Second, it must be possible to choose Δt sufficiently large that the changes in the population over successive timeslices are uncorrelated. (For general Δt this will not be the case. The movements of ancestral lineages in one time step may be correlated with their movements in previous steps if, for example, individuals tend to disperse away from temporarily crowded clusters.) Over all but very small scales, the resulting probability of identity can be written as a function of three parameters: the *effective dispersal rate*, the *neighbourhood size* and the *local scale*. However the usefulness of this result is limited due to a lack of explicit models for which the assumptions can be validated and the effective parameters established. Moreover, as explained in [**BDE02**], although one can in principle extend the formula to approximate the distribution of genealogies amongst larger samples of well-separated genes, additional assumptions need to be made if such genealogies are to be dominated by pairwise coalescence. If several genes are sampled from one location and neighbourhood size is small then multiple coalescence (by which we mean simultaneous coalescence of *three* or more lineages) could become significant.

Multiple merger coalescents have received considerable attention from mathematicians over the last decade. Pitman in [**Pit99**] and Sagitov in [**Sag99**] introduced what we now call Λ-*coalescents*, in which more than two ancestral lineages can coalesce in a single event, but *simultaneous* coalescence events are not allowed. Like Kingman's coalescent, these processes take their values among partitions of $\mathbb{N}$ and their laws can be prescribed by specifying the restriction to partitions of $\{1, 2, \ldots, n\}$ for each $n \in \mathbb{N}$. For our purposes, the Λ-coalescent describes the ancestry of a population whose individuals are labelled by $\mathbb{N}$. Each block in the partition at time t corresponds to a single ancestor at time t before the present, with the elements of the block being the descendants of that ancestor. Tracing backwards in time, the evolution of the Λ-coalescent is as follows: if there are currently p ancestral lineages, then each transition involving j of the blocks merging into one happens at rate

$$\beta^{\Lambda}_{p,j} = \int_{[0,1]} u^{j-2}(1-u)^{p-j}\Lambda(du), \tag{5.2}$$

and these are the only possible transitions. Here, Λ is a finite measure on $[0, 1]$. Kingman's coalescent corresponds to the special case $\Lambda = \delta_0$, the point mass at the origin.

REMARK 5.1. *More generally, one can consider processes with simultaneous multiple coalescence events. Such coalescents were obtained as the genealogies of suitably rescaled population models in* [**MS01**]. *Independently, Schweinsberg in* [**Sch00**] *obtained the same class of coalescents and characterised the possible rates of mergers in terms of a single measure* Ξ *on an infinite simplex. Coalescents which allow* simultaneous *multiple mergers are now generally referred to as* Ξ-coalescents.

Kingman's coalescent can be thought of as describing the genealogy of a random sample from a Fleming-Viot process. In the same way, a Λ-coalescent describes the genealogy of a random sample from a generalised Fleming-Viot process. This process takes its values among probability measures on $[0, 1]$. We will describe it in terms of its generator, $\mathcal{R}$ acting on functions of the form

$$F(\rho) = \int f(x_1, \ldots, x_p)\rho(dx_p)\ldots\rho(dx_1),$$

where $p \in \mathbb{N}$ and $f : [0,1]^p \to \mathbb{R}$ is measurable and bounded. First we need some notation. If $x = (x_1, \dots, x_p) \in [0,1]^p$ and $J \subseteq \{1, \dots, p\}$ we write

$$x_i^J = x_{\min J} \text{ if } i \in J, \text{ and } x_i^J = x_i \text{ if } i \notin J, \qquad i = 1, \dots, p.$$

Then for Λ a finite measure on $[0,1]$, a Λ-Fleming-Viot process has generator

$$\mathcal{R}F(\rho) = \sum_{J \subseteq \{1,\dots,p\}, |J| \geq 2} \beta^{\Lambda}_{p,|J|} \int \big(f(x_1^J, \dots, x_p^J) - f(x_1, \dots, x_p)\big)\, \rho(dx_p) \dots \rho(dx_1),$$

where $\beta^{\Lambda}_{p,j}$ is defined in Equation (5.2). When $\Lambda(\{0\}) = 0$, this can also be written

$$\mathcal{R}F(\rho) = \int_{(0,1]} \int_{[0,1]} \big(F\big((1-u)\rho + u\delta_k\big) - F(\rho)\big)\rho(dk) u^{-2} \Lambda(du).$$

In this case, an intuitive way to think about the process is to consider a Poisson point process on $\mathbb{R}_+ \times (0,1]$ with intensity measure $dt \otimes u^{-2}\Lambda(du)$, which picks jump times and sizes for $\rho(t)$. At a jump time t with corresponding jump size u, a type k is chosen according to $\rho(t-)$, an atom of mass u is inserted at k and $\rho(t-)$ is scaled down by $(1-u)$ so that the total mass remains equal to one, i.e.

$$\rho(t) = (1-u)\rho(t-) + u\delta_k. \tag{5.3}$$

The duality between Λ-coalescents and Λ-Fleming-Viot processes was first proved in [**BL03**]. Their approach uses a correspondence between the Λ-coalescents and stochastic flows of bridges. The duality can also be understood via the Donnelly & Kurtz 'modified lookdown construction' introduced in [**DK99**], and indeed is implicit there. An explicit explanation can be found in [**B&al05**].

In recent work (described briefly in [**Eth08**]), Barton & Etheridge have proposed a new class of consistent forwards and backwards in time models for the evolution of allele frequencies in a population distributed in a two-dimensional (or indeed d-dimensional) spatial continuum which, in the simplest setting, can be thought of as spatial versions of the Λ-Fleming-Viot and Λ-coalescent models (although we emphasize that these are not the same as the spatial Λ-coalescents considered in [**LS06**]). They share many of the advantages of the classical models for spatially structured populations while overcoming at least some of the disadvantages. The idea is simple. Just as in the Λ-Fleming-Viot process, reproduction events are determined by a Poisson point process but now, in addition to specifying a time and a value u, this process prescribes a region of space which will be affected by the event. In what follows the region will be a ball with random centre and radius. Within that region the effect is entirely analogous to Equation (5.3).

This approach differs from existing spatial models in three key ways. First, density dependent reproduction is achieved by basing reproduction events on neighbourhoods (whose locations are determined by the Poisson point process), rather than on individuals. Second, the offspring of a single individual can form a significant proportion of the population in a neighbourhood about the parent, capturing the essentially finite nature of the local population size. Third, large scale extinction-recolonisation events are explicitly incorporated. This reflects the large scale fluctuations experienced by real populations in which the movement and reproductive success of many individuals are correlated. For example, climate change has caused extreme extinction and recolonisation events that dominate the demographic history of humans and other species (e.g. [**EHR04**]).

The spatial Λ-Fleming-Viot process, like its classical counterpart, can be obtained as a limit of individual based models. Those prelimiting models are discussed in [**BEH09**]. In the (backwards in time) spatial Λ-coalescent, ancestral lineages move around according to (dependent) Lévy processes (in fact they will be compound Poisson processes), jumping whenever they are affected by a reproduction event. Two or more lineages can coalesce if they are all affected by the same reproduction event.

Our first aim here is to provide a precise mathematical description of the spatial Λ-Fleming-Viot process and the corresponding spatial Λ-coalescent model and address questions of existence and uniqueness. This is achieved through adapting the work of Evans in [**Eva97**]. The idea is to first construct the dual (backwards in time) process of coalescing Lévy processes corresponding to a finite sample from the population at time zero, and then to use a functional duality to define the forwards in time model. The principal difference between our setting and that of Evans is that, in his work, ancestral lineages evolve *independently* until they meet.

The system of coalescing Lévy processes that describes the genealogy of a sample from the population, mirrors the system of coalescing random walks that plays the same rôle for the stepping stone model. For systems of coalescing walks a number of studies have investigated conditions under which, when viewed on an appropriate timescale, and for sufficiently well-separated samples, the effect of the geographical structure of the population can be summarised as a single 'effective' parameter and the system of coalescing lineages converges to Kingman's coalescent. The first work along these lines appears in [**Cox89**], where Cox considers random walks on a torus $\mathbb{T}(L) \cap \mathbb{Z}^d$ of side length L with the walks coalescing instantly on meeting. This corresponds to taking $G = \mathbb{T}(L) \cap \mathbb{Z}^d$ and $\gamma = \infty$ in Equation (5.1). He shows that if one starts walks from any finite number $n \in \mathbb{N}$ of points chosen independently and uniformly at random from $\mathbb{T}(L) \cap \mathbb{Z}^d$, then in suitable time units, as $L \to \infty$, the number of surviving lineages is determined by Kingman's coalescent. For two spatial dimensions, this analysis was extended by Cox & Durrett in [**CD02**] and Zähle et al. in [**ZCD05**] to random walks on $\mathbb{T}(L) \cap \mathbb{Z}^2$ with delayed coalescence (corresponding to $\gamma < \infty$). It is natural to ask whether similar results are true here. Our second aim then is to establish conditions under which the genealogy of a sample taken at random from a large torus will converge to a non-spatial coalescent. We shall concentrate on the most difficult, but also most biologically relevant, case of two spatial dimensions. If reproduction events only affect bounded neighbourhoods, then, not surprisingly, we recover a Kingman coalescent limit. However, we also consider the more general situation in which in addition to 'small' events that affect only bounded neighbourhoods we allow 'large' extinction-recolonisation events. Unless these events affect a non-negligible proportion of the torus, on a suitable timescale, asymptotically we once again recover a Kingman coalescent. The timescale is determined by the relative rates of 'large' and 'small' events. However, if we have extinction-recolonisation events that affect regions with sidelength of order $\mathcal{O}(L)$, then, again depending on the relative rates of 'large' and 'small' events, we can obtain a more general (non-spatial) Λ-*coalescent* limit or a system of coalescing Brownian motions (where the coalescence is non-local).

The rest of the chapter is laid out as follows. In Section 2 we define the model. In Section 3, we give a precise statement of the conditions under which we obtain convergence of the genealogy of a random sample from a (two-dimensional) torus of side L as $L \to \infty$. The corresponding convergence results are Theorem 5.11 and Theorem 5.15. In Section 4, we establish existence of the process and prove uniqueness in law. In Section 5, we gather the necessary results on Lévy processes in preparation for our proofs of Theorem 5.11 and Theorem 5.15 in Sections 6 and 7.

2. The model

First we describe a prelimiting model. Individuals in our population are assumed to have a *type* taken from $[0, 1]$ and a spatial position in a metric space E that we shall usually take to be $\mathbb{R}^2$ (or $\mathbb{T}(L)$). Even though it will be clear that existence and uniqueness of the process holds in much greater generality, the model is primarily motivated by considerations for populations evolving in two-dimensional continua. The dynamics are driven by a Poisson point process Π on $\mathbb{R}_+ \times \mathbb{R}^2 \times (0,\infty)$ with intensity $dt \otimes dx \otimes \mu(dr)$. If $(t,x,r) \in \Pi$, the first component represents the time of a reproduction event. The event will affect only individuals in $B(x,r)$, the closed ball of centre x and radius r. We require two more ingredients. The first, m, is a fixed positive constant which we shall refer to as the *intensity* of the model. Second, associated to each fixed radius $r > 0$ there is a probability measure ν_r on $[0,1]$. In the sequel, we assume that the mapping $r \mapsto \nu_r$ is measurable with respect to μ.

For definiteness, suppose that the population is initially distributed according to a spatially homogeneous Poisson process. The dynamics of our prelimiting model are described as follows. Suppose that $(t,x,r) \in \Pi$. Consider the population in $B(x,r)$ at time $t-$. If the ball is empty, then nothing happens. Otherwise:

1. Select a 'parent' uniformly at random from those individuals in $B(x,r)$ at time $t-$ and sample $u \in [0,1]$ at random according to ν_r.
2. Each individual in $B(x,r)$, independently, dies with probability u, otherwise it is unaffected by the reproduction event.
3. Throw down offspring in the ball, with the same type as the selected parent (who may now be dead), according to an independent Poisson point process with intensity $um\times \text{Leb}\Big|_{B(x,r)}$ where Leb denotes Lebesgue measure.

We shall refer to these events as *reproduction events*, even though they are also used to model large-scale extinction-recolonisation events. Notice that recolonisation is modelled as being instantaneous even after a large scale extinction.

REMARK 5.2. *For simplicity we have described only a special version of the model in which, even when the reproduction event affects a large region, recolonisation is through a single founder. This guarantees that if we look at the genealogy of a sample from this population, although we may see more than two lineages coalescing in a single event, we do not see* simultaneous *mergers. More generally it would be natural to take a random number of colonists and then, on passing to the limit, the corresponding model would yield a spatial* Ξ*-coalescent.*

Any reproductive event has positive probability of leaving the corresponding region empty, but because the neighbourhoods determined by different reproduction events overlap, an empty region can subsequently become recolonised. Provided the measure $\mu(dr)$ decays sufficiently quickly as $r \to \infty$, Berestycki et al. show in [**BEH09**] that there is a critical value of m above which the population, when started from a translation invariant initial condition, survives with probability one. The difficulty is that it is not easy to find an explicit expression for the distribution of the genealogical trees relating individuals in a sample from the population. Knowing that an ancestral lineage is in a given region of space gives us information about the rate at which that region was hit by reproduction events as we trace backwards in time. On the other hand, simulations reveal that this effect is rarely significant. Mathematically, we overcome this difficulty by considering a model in which the intensity m is infinite, but we preserve some of the signature of a finite local population size by retaining the reproduction mechanism so that a non-trivial proportion of individuals in a neighbourhood are descended from a common ancestor. In particular, this will result in multiple coalescences of ancestral lineages.

Now let us describe the model that arises from letting $m \to \infty$. (That the prelimiting model really does converge to this limit will be proved elsewhere.) At each point $x \in \mathbb{R}^2$, the model specifies a probability measure on type space which we shall write $\rho(t,x,\cdot)$, or sometimes for brevity ρ_x. The interpretation is that if we sample an individual from x, then its type will be determined by sampling from ρ_x. The reproduction mechanism mirrors that for our discrete time model:

DEFINITION 5.3. *The* spatial Λ-Fleming-Viot process, $\{\rho(t,x,\cdot), x \in \mathbb{R}^2, t \geq 0\}$ *specifies a probability measure on the type space* $[0,1]$ *for every* $t \geq 0$ *and every* $x \in \mathbb{R}^2$. *With the notation above, the dynamics of the process are as follows. At every point* (t,x,r) *of the Poisson point process* Π, *we choose* $u \in [0,1]$ *independently according to the measure* $\nu_r(du)$. *We also select a point* z *at random from* $B(x,r)$ *and a type* k *at random according to* $\rho(t-,z,\cdot)$. *For all* $y \in B(x,r)$,

$$\rho(t,y,\cdot) = (1-u)\rho(t-,y,\cdot) + u\delta_k.$$

Sites outside $B(x,r)$ *are not affected, that is* $\rho(t,y,\cdot) = \rho(t-,y,\cdot)$ *for every* $y \notin B(x,r)$.

REMARK 5.4. *There are many variants of this model, some of which are outlined in* **[Eth08]**. *The model presented here should be regarded as fitting into a general framework in which the key feature is that reproduction events are driven by a Poisson point process determining their times and spatial locations, rather than on individuals. Barton et al. in* **[BKE09]** *investigate a version of the model in which, instead of replacing a portion* u *of the population in a disc at the time of a reproduction event, the proportion of individuals affected decays (in a Gaussian distribution) with the distance from the 'centre'* x *of the event. Whereas in the disc based approach in the prelimiting (individual based) model we had to suppress reproduction events that affected empty regions, this is not necessary in the Gaussian model. Moreover, in that setting the prelimiting model has the Poisson point process in* $\mathbb{R}^2$ *with constant intensity* m *as a stationary distribution. Although the proofs would be rather involved, analogues of our results here should carry over to the Gaussian setting.*

Of course we must impose restrictions on the intensity measure if our process is to exist. To see what these should be, consider first the evolution of the probability measure $\rho(t,x,\cdot)$ defining the distribution of types at the point x. This measure experiences a jump of size $y \in A \subseteq (0,1]$ at rate

$$\int_{(0,\infty)} \int_A \pi r^2 \nu_r(du)\mu(dr).$$

By analogy with the Λ-Fleming-Viot process, we expect to require that

$$\Lambda(du) = \int_{(0,\infty)} u^2 r^2 \nu_r(du)\mu(dr) \tag{5.4}$$

defines a finite measure on $[0,1]$. In fact, in the spatial setting we require a bit more. To see why, suppose that ψ is a bounded measurable function on $[0,1]$ and consider the form that the infinitesimal generator of the process must take on test functions of the form $\langle \rho(x,dk), \psi(k)\rangle$ (with angle brackets denoting integration). Denoting the generator (if it exists) by G we will have

$$\begin{aligned} G(\langle\rho,\psi\rangle) &= \int_{\mathbb{R}^2}\int_{(0,\infty)}\int_{[0,1]}\int_{[0,1]} \frac{L_r(x,y)}{\pi r^2}\big(\langle(1-u)\rho(x,\cdot)+u\delta_k,\psi\rangle - \langle\rho(x,\cdot),\psi\rangle\big) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \rho(y,dk)\nu_r(du)\mu(dr)dy \\ &= \int_{\mathbb{R}^2}\int_{(0,\infty)}\int_{[0,1]} \frac{L_r(x,y)}{\pi r^2}\, u\big(\langle\rho(y,\cdot),\psi\rangle - \langle\rho(x,\cdot),\psi\rangle\big)\nu_r(du)\mu(dr)dy, \end{aligned}$$

where $L_r(x,y)$ denotes the volume of the set $B(x,r)\cap B(y,r)$. Notice in particular that $L_r(x,y) \leq \pi r^2 \mathbb{I}_{\{|x-y|\leq 2r\}}$. In the non-spatial case, this term vanishes (set $y = x$), but here if we want the generator to be well-defined on these test functions we make the stronger

ASSUMPTION 5.5.

$$\tilde{\Lambda}(du) = \int_{(0,\infty)} ur^2 \nu_r(du)\mu(dr) \tag{5.5}$$

defines a finite measure on $[0,1]$.

Condition (5.5) controls the jumps of ρ at a single point. Since we are going to follow **[Eva97]** in constructing our process via the dual process of coalescing lineages ancestral to a sample from the population, we should check that such a process is well-defined. First we define the coalescent process more carefully.

In order to make sense of the genealogy of a sample at any time, we extend the Poisson point process Π of reproduction events to the whole time line $(-\infty,+\infty)$. We need some notation for (labelled) partitions.

NOTATION 5.6. (Notation for partitions)

1. *For each integer* $n \geq 1$*, let* $\mathcal{P}_n$ *denote the set of partitions of* $\{1,\dots,n\}$*, and define a labelled partition of* $\{1,\dots,n\}$*, with labels from a set* E*, to be a set of the form* $\{(\pi_1, x_{\pi_1}), \dots, (\pi_k, x_{\pi_k})\}$*, where* $\{\pi_1,\dots,\pi_k\} \in \mathcal{P}_n$ *and* $(x_{\pi_1},\dots,x_{\pi_k}) \in E^k$*. Let* $\mathcal{P}_n^\ell$ *be the set of all labelled partitions of* $\{1,\dots,n\}$*.*
2. *For each* $n \in \mathbb{N}$*, let* $\wp_n$ *denote the partition of* $\{1,\dots,n\}$ *into singletons. Moreover, if* E *is the space of labels and* $\mathbf{x} := (x_1,\dots,x_n) \in E^n$*, let* $\wp_n(\mathbf{x})$ *denote the element* $\{(\{1\},x_1),\dots,(\{n\},x_n)\}$ *of* $\mathcal{P}_n^\ell$*.*
3. *If* $\pi \in \mathcal{P}_n^\ell$ *for some* $n \in \mathbb{N}$*, then* $\mathrm{bl}(\pi)$ *will refer to the unlabelled partition of* $\{1,\dots,n\}$ *induced by* π *and if* $a \in \mathrm{bl}(\pi)$*,* x_a *will be our notation for the label of* a*.*

Our genealogical process will be a labelled partition. As in classical representations of genealogical processes, a block of the partition at genealogical time $t \geq 0$ contains the indices of the initial lineages which share a common ancestor t units of time in the past, and its label gives the current location of this ancestor in $E = \mathbb{R}^2$.

From the description of the forwards-in-time dynamics, the evolution of a sample of ancestral lineages represented by a labelled partition should be the following. We start with a finite collection of lineages at time 0. At each point $(-t,x,r) \in \Pi$ (with $t \geq 0$ here, since genealogical time points towards the past), given that $u \in [0,1]$ is the result of the sampling according to ν_r each lineage present in the ball $B(x,r)$, independently, is affected (resp., is not affected) with probability u (resp., $1-u$). A site y is chosen uniformly in $B(x,r)$, and the blocks of all affected lineages merge into a single block labelled by y. The other blocks and their labels are not modified. We write $\{\mathcal{A}(t),\ t \geq 0\}$ for the Markov process of coalescing lineages described in this way. Its state space is $\bigcup_{n\geq 1} \mathcal{P}_n^\ell$. Note that $\mathcal{A}$ is constructed on the same probability space as that of the Poisson point process of reproduction events. Writing $\mathbb{P}$ for the probability measure on that space, we abuse notation slightly by writing $\mathbb{P}_A$ to indicate that $\mathcal{A}(0) = A$, $\mathbb{P}_A$-a.s. Now let us verify that our condition (5.5) is sufficient to ensure that the process $\{\mathcal{A}(t), t \geq 0\}$ is well-defined. Since two lineages currently at separation $y \in \mathbb{R}^2$ will coalesce if they are *both* involved in a replacement event, which happens at instantaneous rate

$$\int_{(|y|/2,\infty)} L_r(y,0)\left(\int_{[0,1]} u^2\nu_r(du)\right)\mu(dr), \tag{5.6}$$

Condition (5.5) is more than enough to bound the rate of coalescence of ancestral lineages. To guarantee that we can fit together the measures ρ at different points in a consistent way, we also need to be able to control the spatial motion of ancestral lineages. Consider the (backwards in time) dynamics of a single ancestral lineage. It evolves in a series of jumps with intensity

$$dt \otimes \int_{(|x|/2,\infty)} \int_{[0,1]} \frac{L_r(x,0)}{\pi r^2}\, u\, \nu_r(du)\mu(dr)dx \tag{5.7}$$

on $\mathbb{R}_+ \times \mathbb{R}^2$. If we want this to give a well-defined Lévy process, then we require

$$\int_{\mathbb{R}^2} (1 \wedge |x|^2) \left(\int_{(|x|/2,\infty)} \int_{[0,1]} \frac{L_r(x,0)}{\pi r^2}\, u\, \nu_r(du)\mu(dr) \right) dx < \infty. \tag{5.8}$$

But Condition (5.5) certainly guarantees this. In fact it ensures that the rate of jumps of each ancestral lineage is *finite*. In other words ancestral lineages follow compound Poisson processes.

REMARK 5.7. *At first sight it is disappointing that we have to take Condition (5.5) and hence obtain a system of coalescing compound Poisson processes rather than more general symmetric Lévy processes that (5.4) and (5.8) would allow. However, biologically there is not much loss. The 'gap' between Condition (5.5) and the weaker Condition (5.4) is that the latter would allow one to include very large numbers of extremely small jumps (in which only a tiny proportion of the population is affected) as the radius of the area affected by a reproduction event tends to zero. But in our population model, for small r we expect that a* large *proportion of the population in the neighbourhood should be replaced.*

REMARK 5.8. *Notice that the locations of ancestral lineages are* not *independent of one another. Knowing that one lineage has jumped tells us that a reproduction event has taken place that could have affected other lineages ancestral to our sample. In* [**WW02**], *Wilkins & Wakeley consider a somewhat analogous model in which a linear population evolves in discrete generations (see* [**Wil04**] *for a two-dimensional analogue). Each individual in the parental generation scatters an infinite pool of gametes in a Gaussian distribution about themselves, and the next generation is formed by sampling from the pool of gametes at each point. Individuals are assumed to have a finite linear width to avoid the pathologies that arise when common ancestry in a continuum model requires two ancestral lineages to have a physical separation of zero. They observe that 'conditional on not coalescing in the previous generation, two lineages are slightly more likely to be further apart than closer together'. In their setting a change of coordinates settles the problem: the distance apart and the average position of two lineages do evolve independently. For us the dependencies between lineages are more complex because the presence of a jump contains the information that a reproduction event has taken place, whereas the conditioning obviously tells us nothing about the timing of events in the discrete generation model.*

3. The genealogy of points sampled uniformly from a large torus

We now turn our attention to populations evolving on a two-dimensional torus of side L. Our goal is to describe the genealogy of a finite number of individuals sampled uniformly at random from the torus as $L \to \infty$.

To this end, we now consider a family of models indexed by $\mathbb{N}$. For each $L \in \mathbb{N}$, we consider a population evolving on the torus $\mathbb{T}(L) \subset \mathbb{R}^2$ of side L. We identify $\mathbb{T}(L)$ with the subset $[-L/2, L/2]^2$ of $\mathbb{R}^2$ and use the Euclidean norm $|\cdot|$ induced on $\mathbb{T}(L)$ by this identification. Although $B_{\mathbb{T}(L)}(x,r)$ will be our notation for the ball in $\mathbb{T}(L)$ centered in x and with radius r, we shall omit the subscript when there is no risk of confusion.

The population will be subject to two different classes of events that we call 'small' and 'large'. The region affected by each small event will be uniformly bounded (independently of the size of the torus). Large events will affect regions whose diameter is on the order of $\psi(L)$ which will be taken to grow with L, but they will be less frequent. We shall assume that the rate at which a given ancestral lineage is affected by a large event is proportional to $1/\rho(L)$ with $\rho(L)$ also chosen to grow with L.

Now let us make the model more precise. Let $\{\psi(L)\}_{L\geq 1}$ be an increasing sequence such that there exists $\alpha \in (0,1]$ satisfying

$$\lim_{L\to\infty} \frac{\log \psi(L)}{\log L} = \alpha, \tag{5.9}$$

and assume that $|\alpha \log L - \log \psi(L)| = \mathcal{O}((\log L)^{-1/2})$ as $L \to \infty$.

REMARK 5.9. *The latter assumption is not necessary since all our results would still hold with each occurrence of* $(1-\alpha)\log L$ *replaced by* $\log(L\psi(L)^{-1})$ *(see the end of the proof of Proposition 5.26), but it is weak and considerably simplifies the presentation.*

Let $\{\rho(L)\}_{L\geq 1}$ be an increasing sequence with values in $(0,+\infty]$, tending to infinity as $L \to \infty$. Finally, let $\mu^s(dr)$ and $\mu^B(dr)$ be two σ-finite Borel measures on $(0,\infty)$, independent of L, such that there exist some positive constants R^s and R^B satisfying

$$\inf\big\{R : \mu^s\big((R,\infty)\big) = 0\big\} = R^s < \infty \quad \text{and} \quad \inf\big\{R : \mu^B\big((R,\infty)\big) = 0\big\} = R^B < \infty.$$

(For convenience, we ask that $R^B \leq \sqrt{2}/2$ if $\alpha = 1$.) To every $r \geq 0$, we associate two probability measures $\nu_r^s(du)$ and $\nu_r^B(du)$ on $[0,1]$, and we assume that for $\star \in \{B,s\}$ and for each $\varepsilon \in (0,R^\star)$,

$$\mu^\star\big(\big\{r \in [R^\star - \varepsilon, R^\star] : \nu_r^\star \neq \delta_0\big\}\big) > 0. \tag{5.10}$$

If Condition (5.10) does not hold, we decrease the corresponding radius $R^\star$ since otherwise the largest events never affect a lineage.

Let us suppose that for each $L \geq 1$, the reproduction events of the forwards-in-time model can be of two types :

- **Small events**, given by a Poisson point process Π_L^s on $\mathbb{R} \times \mathbb{T}(L) \times (0,\infty)$ with intensity measure $dt \otimes dx \otimes \mu^s(dr)$. If (t,x,r) is a point of Π_L^s, then the centre of the reproduction event is x, its radius is r and the fraction of individuals replaced during the event is chosen according to ν_r^s.
- **Large events**, given by a Poisson point process Π_L^B on $\mathbb{R} \times \psi(L)^{-1}\mathbb{T}(L) \times (0,\infty)$, independent of Π_L^s and with intensity measure $\rho(L)^{-1}dt \otimes dx \otimes \mu^B(dr)$. If (t,x,r) is a point of Π_L^B, then the centre of the reproduction event is $\psi(L)x \in \mathbb{T}(L)$, its radius is $\psi(L)r$ and the fraction of individuals replaced during the event is chosen according to ν_r^B.

Notice that we allow $\rho(L)$ to be infinite, in which case large events do not occur. Of course, this could be formulated in terms of a single Poisson point process, but our aim here is to disentangle the effects of events of different scales.

REMARK 5.10. *The formulation adopted to introduce large events emphasizes the fact that a single lineage is affected by a large event at a rate proportional to* $\rho(L)^{-1}$. *It is equivalent to saying that large events correspond to a Poisson point process on* $\mathbb{R}_+ \times \mathbb{T}(L) \times (0,\infty)$ *with intensity* $(\rho(L)\psi(L)^2)^{-1}dt \otimes dx \otimes \mu^B(dr)$. *At a point* (t,x,r) *of this process, a reproduction event occurs in which a portion* $u \sim \nu_r^B$ *of the population in* $B(x,\psi(L)r)$ *is replaced.*

In order for the genealogical processes, which we now denote by $\mathcal{A}^L$ to emphasize dependence on L, to be well-defined for every $L \in \mathbb{N}$, we assume that Condition (5.5) is fulfilled. In this setting, the condition can be written

$$\int_0^{R^s}\int_0^1 r^2u\,\nu_r^s(du)\mu^s(dr)+\frac{1}{\rho(L)}\int_0^{R^B}\int_0^1 r^2u\,\nu_r^B(du)\mu^B(dr)<\infty.$$

Let us introduce some more notation. We write

$$\Gamma(L,1)\equiv\left\{x\in\mathbb{T}(L):|x|\geq\frac{L}{\log L}\right\},$$

and for each integer $n\geq 2$,

$$\Gamma(L,n):=\left\{\{x_1,\dots,x_n\}\in\mathbb{T}(L)^n:|x_i-x_j|\geq\frac{L}{\log L}\ \text{ for all } i\neq j\right\},$$
$$\Gamma_{\mathcal{A}}(L,n):=\left\{\{(a_1,x_{a_1}),\dots,(a_k,x_{a_k})\}\in\mathcal{P}_n^\ell:\ \{x_{a_1},\dots,x_{a_k}\}\in\Gamma(L,k)\right\},$$

where as before $\mathcal{P}_n^\ell$ denotes the labelled partitions of $\{1,\dots,n\}$. When we require an element A of $\Gamma_{\mathcal{A}}(L,n)$ to have exactly n blocks, we shall write $A\in\Gamma_{\mathcal{A}}(L,n)^*$.

In order to obtain a non-trivial limit, we rescale time for the process $\mathcal{A}^L$ by a factor that we denote ϖ_L. Recall that if $A\in\mathcal{P}_n^\ell$ for some $n\in\mathbb{N}$, $\mathrm{bl}(A)$ stands for the unlabelled partition of $\{1,\dots,n\}$ induced by A. For each $L\in\mathbb{N}$, let us define the (non-Markov) process $\mathcal{A}^{L,u}$ by

$$\mathcal{A}^{L,u}(t)=\mathrm{bl}\big(\mathcal{A}^L(\varpi_L t)\big),\qquad t\geq 0.$$

Note that for each $L\in\mathbb{N}$, if we start $\mathcal{A}^L$ from A_L, a labelled partition of $\{1,\dots,n\}$ with labels from $\mathbb{T}(L)$, then $\mathcal{A}^{L,u}$ takes its values in the Skorohod space $D_{\mathcal{P}_n}[0,\infty)$ of all càdlàg paths with values in $\mathcal{P}_n$ (the set of partitions of $\{1,\dots,n\}$), $\mathbb{P}_{A_L}$-a.s.

Recall the definition of α given in (5.9). In the absence of large events, our model is similar in many respects to the two-dimensional stepping stone model and so it comes as no surprise that just as for the stepping stone model studied, the genealogy of a random sample from the torus should converge (on a suitable timescale) to a Kingman coalescent as the size of the torus tends to infinity (see in particular [**CG86, CG90**] and [**CD02, ZCD05**] for precise statements of this result in different contexts). Our first result says that if $\alpha<1$, then we still obtain a Kingman coalescent, but the *timescale* will be influenced by the large events: the effect of the latter is to reduce the effective population size.

Before stating the result formally, let us try to understand why we should expect something like this to be true. To understand the appropriate timescale we just need to consider two lineages. The time they need to coalesce will be decomposed into two phases. If $\rho(L)$ is not too big, the first phase will be the time until they first come within distance $2R^B\psi(L)$ and the second will be the additional time required for them to coalesce. During the first phase they evolve according to independent compound Poisson processes. If $\rho(L)$ is small enough, the coalescence event awaited during the second phase will, with probability close to one, be triggered by a large event. For larger values of $\rho(L)$, large events will not be frequent enough to hit the two lineages when they are at a distance that would allow them to coalesce (i.e., less than $2R^B\psi(L)$), and coalescence will instead be caused by a small-scale event. The first phase is then taken to be the time until the lineages first come within distance $2R^s$ of one another. The fact that with high probability they won't be hit by the same large-scale event means that once again they evolve (almost) independently of one another during this first phase. The second phase is now the time taken for them to coalesce due to a small event. The transition between these two regimes is when $\rho(L)\propto\psi(L)^2\log L$. Now suppose that we start from a sample in $\Gamma(L,n)$. The first phase is then long enough that, when it ends, the spatial location of lineages is no longer correlated with their starting points. Finally, why do large-scale events not lead to multiple mergers? The key point is that, when a pair of lineages ancestral to our sample first comes within $2R^B\psi(L)$ of one another, all *other* pairs are still well-separated. So if $\rho(L)$ is not too big, this pair will coalesce before a third lineage can come close enough to be affected by a common event. If we take larger $\rho(L)$, the reason is exactly the same but now lineages have to come within distance $2R^s$ and coalescence is driven by small events.

Here then is the formal result. In the following, σ_s^2 (resp., $\sigma_B^2\psi(L)^2$) is the variance of the displacement of a lineage during one unit of time due to small (resp., large) events, see (5.23) below.

THEOREM 5.11. *Let $\mathcal{K}$ denote Kingman's coalescent, and recall that for each $n \in \mathbb{N}$, $\wp_n$ denotes the partition of $\{1, \ldots, n\}$ into singletons. In the notation of (5.9), suppose $\alpha < 1$ (and (5.10) holds). Then, for each integer $n \geq 2$ and any sequence $(A_L)_{L\in\mathbb{N}}$ such that $A_L \in \Gamma_{\mathcal{A}}(L, n)^*$ for every L,*

$$\mathcal{L}_{\mathbb{P}_{A_L}}(\mathcal{A}^{L,u}) \Rightarrow \mathcal{L}_{\mathbb{P}_{\wp_n}}(\mathcal{K}) \qquad \text{as } L \to \infty,$$

where

$$\varpi_L = \begin{cases} \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2} & \text{if } \rho(L)^{-1}\psi(L)^2 \to \infty, \\ \frac{(1-\alpha)L^2\log L}{2\pi(\sigma_s^2+b\sigma_B^2)} & \text{if } \rho(L)^{-1}\psi(L)^2 \to b \in [0,\infty) \text{ and } \frac{\psi(L)^2\log L}{\rho(L)} \to \infty, \\ \frac{L^2\log L}{2\pi\sigma_s^2} & \text{if } \rho(L)^{-1}\psi(L)^4 \to \mathfrak{b} \in [0,\infty) \text{ or } \frac{L^2\log L}{\rho(L)} \to 0. \end{cases}$$

Here $\mathcal{L}_{\mathrm{P}}(X)$ denotes the law under the probability measure P of the random variable X and $\Rightarrow$ refers to weak convergence of probability measures.

For $\alpha = 1$, things are more complicated. Now large scale events cover a non-negligible fraction of the torus. If they happen too quickly, then they will be able to capture multiple lineages while the locations of those lineages are still correlated with their starting points. For intermediate ranges of $\rho(L)$, lineages will have homogenised their positions on $\mathbb{T}(L)$ through small events, but not coalesced, before the first large event occurs and we can expect a Λ-coalescent limit. If the large events are too rare, then coalescence will be through small events and we shall recover the Kingman coalescent again.

To give a precise result we need to define the limiting objects that arise. In the case $\alpha = 1$, for each $L \in \mathbb{N}$, we set

$$\varpi_L = \begin{cases} \rho(L) & \text{if } \rho(L)/(L^2\log L) \text{ has a finite limit,} \\ \frac{L^2\log L}{2\pi\sigma_s^2} & \text{if } \rho(L)/(L^2\log L) \to +\infty, \end{cases}$$

and define $\mathcal{A}^{L,u}$ as before. Since we shall need to keep track of the labels (spatial positions) of the ancestral lineages in some cases, it will also be convenient to introduce the following rescaling of $\mathcal{A}^L$, evolving on $\mathbb{T}(1)$ for all $L \in \mathbb{N}$:

$$\bar{\mathcal{A}}^L(t) = \frac{1}{L}\,\mathcal{A}^L(\varpi_L t), \qquad t \geq 0,$$

where by this notation we mean that the labels are rescaled by a factor L^{-1}. Similarly, for $\mathbf{x} \in \mathbb{T}(1)^n$ we write $L\mathbf{x}$ for $(Lx_1, \ldots, Lx_n) \in \mathbb{T}(L)^n$. Finally, let us introduce the processes which will appear as the limits of our rescaled genealogical processes.

DEFINITION 5.12. *Let $b \in [0,\infty)$. We call $\bar{\mathcal{A}}^{\infty,b}$ the Markov process with values in $\bigcup_{n\in\mathbb{N}}\mathcal{P}_n^\ell$ (with labels in $\mathbb{T}(1)$) such that*

1. *The labels of the lineages perform independent Brownian motions on $\mathbb{T}(1)$ at speed $b\sigma_s^2$ (if $b = 0$, the labels are constant), until the first large event occurs.*

2. *Large events are generated by a Poisson point process* $\overline{\Pi}^B$ *on* $\mathbb{R}\times\mathbb{T}(1)\times(0,\sqrt{2}/2]$ *with intensity measure* $dt\otimes dx\otimes\mu^B(dr)$. *At a point* $(-t,x,r)$ *of* $\overline{\Pi}^B$, *a number* $u\in[0,1]$ *is sampled from the probability measure* ν_r^B, *and each lineage whose label belongs to* $B_{\mathbb{T}(1)}(x,r)$ *is affected (resp., is not affected) by the event with probability* u *(resp.,* $1-u$*), independently of each other. A label* z *is chosen uniformly at random in* $B_{\mathbb{T}(1)}(x,r)$*, and all the lineages affected merge into one block which adopts the label* z*. The other lineages (blocks and labels) remain unchanged.*
3. *The evolution of the labels starts again in the same manner.*

REMARK 5.13. *Notice that this process looks like another spatial* Λ*-coalescent, except that now ancestral lineages perform independent spatial motions in between coalescence events. This process is dual (in the obvious way) to a spatial* Λ*-Fleming-Viot process in which, during their lifetimes, individuals move around in space according to independent Brownian motions.*

For each $r\in[0,\sqrt{2}/2]$, let V_r denote the volume of the ball $B_{\mathbb{T}(1)}(0,r)$.

DEFINITION 5.14. *Let* $\beta\in[0,\infty)$. *We use* $\Lambda^{(\beta)}$ *to denote the* Λ*-coalescent, defined on* $\bigcup_{n\in\mathbb{N}}\mathcal{P}_n$*, for which if there are currently* m *ancestral blocks, then each transition involving* k *of them merging into one happens at rate*

$$\lambda_{m,k}^{(\beta)}=\int_0^{(\sqrt{2})^{-1}}\int_0^1 (V_r u)^k(1-V_r u)^{m-k}\nu_r^B(du)\mu^B(dr)+\beta\,\delta_{\{k=2\}}.$$

Recall the notation $\wp_n$ and $\wp_n(\mathbf{x})$ introduced in 5.6, and $\mathcal{L}_{\mathrm{P}}(X)$ and $\Rightarrow$ introduced in the statement of Theorem 5.11. We can now state the result for $\alpha=1$.

THEOREM 5.15. *Let* $n\in\mathbb{N}$, $\mathbf{x}\in\mathbb{T}(1)^n$ *such that* $x_i\neq x_j$ *whenever* $i\neq j$*, and let* $(A_L)_{L\in\mathbb{N}}$ *be such that for every* L, $A_L\in\Gamma_{\mathcal{A}}(L,n)^*$. *Then, as* $L\to\infty$*,*
(a) *If* $\rho(L)L^{-2}\to b\in[0,\infty)$*,*

$$\mathcal{L}_{\mathbb{P}_{\wp_n(L\mathbf{x})}}\big(\bar{\mathcal{A}}^L\big)\Rightarrow\mathcal{L}_{\mathbb{P}_{\wp_n(\mathbf{x})}}\big(\bar{\mathcal{A}}^{\infty,b}\big).$$

(b) *If* $\rho(L)L^{-2}\to\infty$, $\frac{2\pi\sigma_s^2\rho(L)}{L^2\log L}\to\beta\in[0,\infty)$ *and if the total rate of occurrence of large events is finite (i.e.,* μ^B *has finite total mass),*

$$\mathcal{L}_{\mathbb{P}_{A_L}}\big(\mathcal{A}^{L,u}\big)\Rightarrow\mathcal{L}_{\mathbb{P}_{\wp_n}}\big(\Lambda^{(\beta)}\big).$$

(c) *If* $\frac{\rho(L)}{L^2\log L}\to\infty$*,*

$$\mathcal{L}_{\mathbb{P}_{A_L}}\big(\mathcal{A}^{L,u}\big)\Rightarrow\mathcal{L}_{\mathbb{P}_{\wp_n}}\big(\mathcal{K}\big).$$

The intuition behind this result is as follows. If $\alpha=1$ large events cover a non-negligible fraction of the torus, and so only a few large events are sufficient to gather two lineages at a distance at which they can coalesce. However, a local central limit theorem will give us that on a timescale of order at most $\mathcal{O}(L^2)$, a lineage subject to only small events behaves approximately like Brownian motion, whereas after a time $t_L\gg L^2$, its distribution is nearly uniform on $\mathbb{T}(L)$ (for L large enough, see Lemma 5.23). Since the mean time before a large event affects a lineage is of order $\mathcal{O}(\rho(L))$, the limiting genealogical process (when we include both large and small reproduction events) will depend on how $\rho(L)$ scales with L^2. If $\rho(L)$ is of order at most $\mathcal{O}(L^2)$, then space matters and the process $\mathcal{A}^L$ rescaled to evolve on $\mathbb{T}(1)$ on the timescale $\rho(L)$ converges to a system of coalescing Brownian motions, whereas if $\rho(L)\gg L^2$, the homogenisation of the labels/locations of the lineages before the occurrence of the first large event which affects them leads to a limiting unlabelled genealogical process given by an exchangeable coalescent with multiple mergers.

REMARK 5.16. *It is somehow disappointing that we must impose a finite rate of large events to obtain the convergence of Theorem 5.15(b). Indeed, it seems that case (a) should give us the right picture: in the limit, in between large events lineages perform Brownian motions on the torus of side 1 due to small events, except that now the time required for at least one lineage to be affected by a large event is so long that lineages exhaust space and their locations become uniformly distributed over the torus before they are taken by a coalescence event. However, when μ^B has infinite mass, lineages are infinitely often in the (geographical) range of a large reproduction event over any interval of time, and we need good control of their complete paths to actually be able to say something about the epoch and outcome of the first potential coalescence event. Now, observe that Equation (5.65) can only be generalized to the finite-dimensional distributions of these paths, and does not guarantee that a large event cannot capture some of the lineages at a time when they are not uniformly distributed over* $\mathbb{T}(1)$.

4. Existence and uniqueness of the forwards-in-time process

Our spatial Λ-Fleming-Viot process associates a probability measure on type space to each point in $\mathbb{R}^2$. In other words, it takes its values among functions from $\mathbb{R}^2$ to $\mathcal{M}_1([0,1])$. In [**Eva97**], Evans uses duality with a system of coalescing Borel right processes on a Lusin space E to construct a family of Markov processes with values in the set of functions from E to $\mathcal{M}_1(\{0,1\}^{\mathbb{N}})$ (or equivalently, to $\mathcal{M}_1([0,1])$). He also obtains uniqueness in distribution of the process. In his setting, coalescing particles evolve independently until they meet, at which point they instantly coalesce. In our case, the particles in the candidate dual do not move independently and nor do two particles hit by the same reproduction event necessarily coalesce, but nonetheless the key ideas from his construction remain valid. Note that, although we present the result in two dimensions, the proof carries over to other dimensions.

First we give a formal description of the coalescing dual and then we use the Evans' construction to give existence and uniqueness in law of a process ρ which assigns a probability measure on $[0,1]$ to each point in $\mathbb{R}^2$. We then identify ρ as the spatial Λ-Fleming-Viot process in which we are interested.

4.1. State-space of the process and construction via duality. We shall only present the main steps of the construction, and refer to [**Eva97**] for more details.

Let us define $\tilde{\Xi}$ as the space of all Lebesgue-measurable maps $\rho : \mathbb{R}^2 \to \mathcal{M}_1([0,1])$. Two elements ρ_1 and ρ_2 of $\tilde{\Xi}$ are said to be equivalent if $\mathrm{Leb}(\{x \in \mathbb{R}^2 : \rho_1(x) \neq \rho_2(x)\}) = 0$. Let Ξ be the quotient space of $\tilde{\Xi}$ by this equivalence relation. If E is a compact space, let us write $C(E)$ for the Banach space of all continuous functions on E, equipped with the supremum norm $\|\cdot\|_\infty$. For each $n \in \mathbb{N}$, let $L^1(C([0,1]^n))$ be the Banach space of all Lebesgue-measurable maps $\Phi : (\mathbb{R}^2)^n \to C([0,1]^n)$ such that $\int_{(\mathbb{R}^2)^n} \|\Phi(x)\|_\infty \, dx < \infty$. A remark in Section 3 of [**Eva97**] tells us that the separability of $L^1(C([0,1]))$ and a functional duality argument guarantee that Ξ, equipped with the relative weak* topology, is a (compact) metrisable space. Finally, if λ is a measure on a space E', let us write $L^1(\lambda)$ for the set of all measurable functions $f : E' \to \mathbb{R}$ such that $\int_{E'} |f(e)|\lambda(de) < \infty$.

Let $n \in \mathbb{N}$. Given $\Phi \in L^1(C([0,1]^n))$, let us define a function $I_n(\cdot\,;\Phi) \in C(\Xi)$ by

$$I_n(\rho;\Phi) := \int_{(\mathbb{R}^2)^n} \Big\langle \bigotimes_{1\leq i\leq n} \rho(x_i), \Phi(x_1,\dots,x_n)\Big\rangle \, dx_1 \dots dx_n,$$

where as before the notation $\langle \nu, f\rangle$ stands for the integral of the function f against the measure ν. We have the following lemma, whose proof is essentially that of Lemma 3.1 in [**Eva97**].

LEMMA 5.17. *The linear subspace spanned by the constant functions and functions of the form* $I_n(\cdot\,;\Phi)$*, with* $\Phi = \psi \otimes \left(\prod_{i=1}^n \chi_i\right)$*,* $\psi \in L^1(dx^{\otimes n}) \cap C((\mathbb{R}^2)^n)$ *and* $\chi_i \in C([0,1])$ *for all* $1 \leq i \leq n$ *is dense in* $C(\Xi)$*.*

We need a last definition before stating the existence and uniqueness result. Let $n \in \mathbb{N}$. For any $\rho \in \Xi$, $\pi \in \mathcal{P}_n^\ell$ such that $\mathrm{bl}(\pi) = \{a_1, \dots, a_k\}$, and any bounded measurable function $F : [0,1]^n \to \mathbb{R}$, we set

$$\Upsilon_n(\rho;\pi;F) := \int_{[0,1]^k} F(v_{a^{-1}(1)}, \dots, v_{a^{-1}(n)})\rho(x_{a_1})(dv_{a_1}) \dots \rho(x_{a_k})(dv_{a_k}),$$

where $a^{-1}(i)$ is the (unique) block a_j which contains i. In words, we assign the same variable to all coordinates which belong to the same block in the partition π. (Recall that x_a is our notation for the label of block a.) Recall also the notation $\wp_n(\mathbf{x})$ and $\mathcal{A}$ introduced in Notation 5.6 and the following paragraph.

THEOREM 5.18. *There exists a unique, Feller, Markov semigroup* $\{Q_t, t \geq 0\}$ *on* Ξ *such that for all* $n \in \mathbb{N}$ *and* $\Phi \in L^1(C([0,1]^n))$*, we have*

$$\int Q_t(\rho, d\rho') I_n(\rho';\Phi) = \int_{(\mathbb{R}^2)^n} \mathbb{E}_{\wp_n(\mathbf{x})}\big[\Upsilon_n\big(\rho;\mathcal{A}(t);\Phi(x_1,\dots,x_n)\big)\big]dx_1 \dots dx_n. \tag{5.11}$$

Consequently, there exists a Hunt process $\{\rho(t), t \geq 0\}$ *with state-space* Ξ *and transition semigroup* $\{Q_t, t \geq 0\}$*.*

Before proving Theorem 5.18, let us make two comments on this result. First, since the Ξ-valued process we obtain is a Hunt process it is quasi-left continuous, that is, it is almost surely left-continuous at any previsible stopping time (see e.g. [**Ber96**] for a definition of quasi-left continuous filtrations). However, if we allow very small events to accumulate in any finite region of $\mathbb{R}^2$, it seems difficult to show that the local fluctuations induced are not strong enough to prevent ρ from being at least càdlàg. The regularity of the spatial Λ-Fleming Viot process is yet an interesting question which will require a thorough investigation.

Second, as in the classical stepping stone model introduced in (5.1), the duality relation (5.11) can be interpreted in terms of genealogies of a sample of individuals. Indeed, recall the relations (1.29) between the stepping stone model and the system $(\{n_i(t);\ i \in I\})_{t \geq 0}$ of particles migrating from deme i to deme j at rate m_{ji} and coalescing by pairs at rate $1/N_e$ when in the same deme. These equations show that a function (here the $n_i(0)$-th moments) of the frequencies at different sites of $\mathbb{Z}^2$ and at (forward) time t can be expressed in terms of the genealogy of a sample made of $n_i(0)$ individuals in deme i for every $i \in I$, and run for a (backward) time t: all lineages having coalesced by time t necessarily carry the same type, whose law is given by the type distribution at the site where their ancestor lies at backward time t (or forward time 0). Equation (5.11) can be interpreted in exactly the same manner, but holds for a much wider collection of functions of ρ and $\mathcal{A}$.

Proof of Theorem 5.18. The observation that the construction of Evans can also be justified in our setting follows from Remark (a) at the end of his Section 4.

First, we must verify consistency of $\mathcal{A}$ in the sense of his Lemma 2.1. In fact this is the 'sampling consistency' described in the introduction and was a primary consideration in writing down our model. It follows since the movement of the labels of a collection of blocks does not depend on the blocks themselves and from the fact that a coalescence event of the form $\{(\{1\}, x_1), (\{2\}, x_2)\} \to \{(\{1,2\}, x)\}$ for a pair of particles corresponds to a jump $\{(\{1\}, x_1)\} \to \{(\{1\}, x)\}$ onto the same site $x \in \mathbb{R}^2$ if we restrict our attention to the first particle.

The next property needed in the construction is that provided it is true at $t = 0$, for every $t > 0$ the distribution of the labels in $\mathcal{A}(t)$ has a Radon-Nikodym derivative with respect to Lebesgue measure, and furthermore an analogue of Evans' Equation (4.2) holds. In the setting of [**Eva97**], the first requirement follows from independence of the spatial motions followed by different labels and the corresponding result for a single label. Here, since the motion of all lineages is driven by the same Poisson process of events, their movements are correlated. However, note that each jump experienced by a lineage in the interval $[-t, 0]$ takes it to a position that is uniformly distributed over the open ball affected by the corresponding reproduction event. Thus, if $\mathcal{A}(t)$ has k blocks and $D \subset (\mathbb{R}^2)^k$ has zero Lebesgue measure, the probability that the labels of the blocks of $\mathcal{A}(t)$ belong to D is equal to 0. Equation (4.2) of [**Eva97**] then still holds, without Evans' additional assumption of the existence of a dual process for the motion of one lineage (which anyway is satisfied since our lineages perform symmetric Lévy processes).

The last step is to check the strong continuity of the semigroup $\{Q_t, t \geq 0\}$, but this readily follows from the relation (5.11) and the Feller property of $\mathcal{A}$ (which is itself evident since jumps do not accumulate in our dual process).

The desired conclusion now follows from Theorem 4.1 in [**Eva97**]. □

4.2. Identification of the process. We can use (5.11) to derive an expression for the infinitesimal generator of $\{\rho(t), t \geq 0\}$ acting on the functions $I_n(\cdot\,;\Phi)$ considered in Lemma 5.17. This lemma and the uniqueness result stated in Theorem 5.18 guarantee that it will be sufficient to characterize the process ρ and to show that it corresponds to the evolution we described in Section 2 in terms of a Poisson point process of reproduction events.

Let $n \in \mathbb{N}$ and $\Phi \in C(\Xi)$ be such that $\Phi = \psi \otimes \big(\prod_{i=1}^n \chi_i\big)$, where $\psi \in L^1(dx^{\otimes n}) \cap C((\mathbb{R}^2)^n)$ and $\chi_i \in C([0,1])$ for all $1 \leq i \leq n$. Writing G the generator of the process ρ and $\mathcal{G}_n$ the generator of the coalescing Lévy processes $\mathcal{A}$ acting on functions of $\mathcal{P}_n^{\ell}$, we obtain from (5.11) that

$$\begin{aligned}
GI_n(\,\rho\,;\Phi) &= \lim_{t\to 0} \frac{\mathbb{E}_\rho[I_n(\rho(t),\Phi)] - I_n(\rho,\Phi)}{t} \\
&= \lim_{t\to 0} \frac{1}{t} \int_{(\mathbb{R}^2)^n} \psi(x_1,\ldots,x_n) \bigg\{ \mathbb{E}_{\wp_n(\mathbf{x})}\Big[\Upsilon_n\Big(\rho; \mathcal{A}(t); \prod_{i=1}^n \chi_i\Big)\Big] - \prod_{i=1}^n \langle \rho(x_i), \chi_i\rangle \bigg\}\, dx^{\otimes n} \\
&= \int_{(\mathbb{R}^2)^n} \psi(x_1,\ldots,x_n)\, \mathcal{G}_n\Big[\Upsilon_n\Big(\rho;\, \cdot\,; \prod_{i=1}^n \chi_i\Big)\Big](\wp_n(\mathbf{x}))\, dx^{\otimes n}. \qquad (5.12)
\end{aligned}$$

Note that the quantity on the right-hand side of (5.12) is well-defined (and the interchange of limit and integral is valid), since ψ belongs to $L^1(dx^{\otimes n})$ and the rate at which at least one of $k \leq n$ blocks is affected by a reproduction event is bounded by n times the integral in (5.5), so that $\mathcal{A}$ is a jump-hold process and its generator satisfies

$$\Big\| \mathcal{G}_n\Big[\Upsilon_n\Big(\rho;\, \cdot\,; \prod_{i=1}^n \chi_i\Big)\Big]\Big\|_\infty \leq 2Cn \,\Big\| \Upsilon_n\Big(\rho;\, \cdot\,; \prod_{i=1}^n \chi_i\Big)\Big\|_\infty \leq 2Cn \prod_{i=1}^n \|\chi_i\|_\infty < \infty$$

for a given constant $C < \infty$.

Using the description of the evolution of $\mathcal{A}$ in terms of events in Π, the right-hand side of (5.12) is equal to

$$\begin{aligned}
&\int_{(\mathbb{R}^2)^n} dx^{\otimes n}\ \psi(x_1,\ldots,x_n) \int_{\mathbb{R}^2} dy \int_0^\infty \mu(dr) \int_0^1 \nu_r(du) \int_{B(y,r)} \frac{dz}{\pi r^2} \\
&\qquad \times \sum_{I \subset \{1,\ldots,n\}} \Big[\prod_{i\in I} \mathbb{I}_{B(y,r)}(x_i) \prod_{i' \notin I} \mathbb{I}_{B(y,r)^c}(x_{i'}) \Big]
\end{aligned}$$

$$\times \sum_{J\subset I} u^{|J|}(1-u)^{|I|-|J|}\Big[\prod_{i\notin J}\langle \rho(x_i),\chi_i\rangle\Big]\Big[\Big\langle \rho(z),\prod_{j\in J}\chi_j\Big\rangle - \prod_{j\in J}\langle\rho(x_j),\chi_j\rangle\Big], \tag{5.13}$$

where $|\cdot|$ stands for cardinality. Indeed, given $x_1,\dots,x_n$ in (5.13), only one term in the sum over $I\subset\{1,\dots,n\}$ is non-zero. For this particular term, each of the $|I|$ blocks whose labels lie in $B(y,r)$ belong to the set J of the blocks affected by the event with probability u (independently of one another), and the affected blocks adopt the label z. After some algebra and several uses of Fubini's theorem, we obtain that (5.13) is equal to

$$\begin{aligned}\int_{\mathbb{R}^2} dy \int_0^\infty \mu\,(dr)\int_0^1 \nu_r(du)\int_{B(y,r)}\frac{dz}{\pi r^2}\int_0^1 \rho_z(dk)\int dx_1\dots dx_n\,\psi(x_1,\dots,x_n)&\\ \times \sum_{I\subset\{1,\dots,n\}}\prod_{j\notin I}\{\mathbb{I}_{B(y,r)^c}(x_j)\langle\rho_{x_j},\chi_j\rangle\}\prod_{i\in I}\mathbb{I}_{B(y,r)}(x_i)&\\ \times\Big(\prod_{i\in I}\langle(1-u)\rho_{x_i}+u\delta_k,\chi_i\rangle-\prod_{i\in I}\langle\rho_{x_i},\chi_i\rangle\Big),&\end{aligned}\tag{5.14}$$

which is precisely the generator of the forwards-in-time process of Section 2. Using Theorem 5.18, we arrive at the following result.

PROPOSITION 5.19. *The martingale problem associated to the operator G defined by (5.14) on functions of the form given in Lemma 5.17 is well-posed. Furthermore, the spatial Λ-Fleming-Viot process ρ of Theorem 5.18 is solution to it.*

5. Some estimates for symmetric Lévy processes

In this section, we gather some results on symmetric Lévy processes that we shall need to call upon in our proofs of Theorem 5.11 and Theorem 5.15. First, we introduce some notation that we shall use repeatedly.

NOTATION 5.20.

1. *In the following, we shall suppose that all the random objects considered are constructed on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$, and if X is a process defined on Ω with state-space E and $x\in E$, we shall write $\mathbb{P}_x$ for the probability measure on Ω under which $X(0)=x$ a.s.*
2. *For a stochastic process $\{X_t\}_{t\geq 0}$ evolving in $\mathbb{T}(L)$, we shall write $T(R,X)$ for the* first entrance time *of X into $B_{\mathbb{T}(L)}(0,R)$. When there is no ambiguity we write simply $T(R)$.*

Let $(\ell^L)_{L\geq 1}$ be a sequence of Lévy processes such that for each $L\in\mathbb{N}$, ℓ^L evolves on the torus $\mathbb{T}(L)$ and $\ell^L(1)-\ell^L(0)$ has a covariance matrix of the form $\sigma_L^2\mathrm{Id}$. Assume that the following conditions hold.

ASSUMPTION 5.21. (i) *There exists $\sigma^2>0$ such that $\sigma_L^2\to\sigma^2$ as $L\to\infty$.*
(ii) $\mathbb{E}_0\big[|\ell^L(1)|^4\big]$ *is bounded uniformly in L.*

Our first lemma describes the time ℓ^L needs to reach a ball of radius $d_L \ll L$ around 0, when it starts at distance $\mathcal{O}(L)$ of the origin.

LEMMA 5.22. *Let $(d_L)_{L\geq 1}$ be such that $\liminf_{L\to\infty} d_L>0$ and $\frac{\log^+(d_L)}{\log L}\to\gamma\in[0,1)$ as $L\to\infty$. Then,*

$$\lim_{L\to\infty}\sup_{t\geq 0}\sup_{x_L\in\Gamma(L,1)}\left|\mathbb{P}_{x_L}\left[T(d_L,\ell^L)>\frac{(1-\gamma)L^2\log L}{\pi\sigma^2}\,t\right]-e^{-t}\right|=0. \tag{5.15}$$

The proof of Lemma 5.22 follows that of Theorem 2 in [**CD02**]. In particular, we shall use the following local central limit theorem (which is the counterpart in our setting of Lemma 3.1 in [**CD02**]). Let $\lfloor z\rfloor$ denote the integer part of $z\in\mathbb{R}$, and write $p^L(x,t)$ for $\mathbb{P}_x[\ell^L(t)\in B(0,d_L)]$.

LEMMA 5.23. (a) *Let* $\varepsilon_L=(\log L)^{-1/2}$. *There exists a constant* $C_1<\infty$ *such that for every* $L\geq 2$,

$$\sup_{t\geq\lfloor\varepsilon_L L^2\rfloor}\ \sup_{x\in\mathbb{T}(L)}\frac{\lfloor\varepsilon_L L^2\rfloor}{d_L^2}\,p^L(x,t)\leq C_1. \tag{5.16}$$

(b) *If* $v_L\to\infty$ *as* $L\to\infty$, *then*

$$\lim_{L\to\infty}\ \sup_{t\geq\lfloor v_L L^2\rfloor}\ \sup_{x\in\mathbb{T}(L)}\frac{L^2}{d_L^2}\left|\,p^L(x,t)-\frac{\pi d_L^2}{L^2}\right|=0. \tag{5.17}$$

(c) *If* $u_L\to\infty$ *as* $L\to\infty$ *and* $I(d_L,x):=1+(|x|^2\vee d_L^2)$, *then*

$$\lim_{L\to\infty}\ \sup_{x\in\mathbb{T}(L)}\ \sup_{u_L I(d_L,x)\leq t\leq\varepsilon_L L^2}\left|\frac{2\sigma_L^2 t}{d_L^2}\,p^L(x,t)-1\right|=0. \tag{5.18}$$

(d) *There exists a constant* $C_2<\infty$ *such that for every* $L\geq 1$,

$$\sup_{t\geq 0}\ \sup_{x\in\mathbb{T}(L)}\left(1+\frac{|x|^2}{d_L^2}\right)p^L(x,t)\leq C_2. \tag{5.19}$$

Since the proofs of both lemmas are highly reminiscent of those of Theorem 2 and Lemma 3.1 in [**CD02**], we shall only give the arguments we need to modify and refer to their paper for more extensive proofs.

Proof of Lemma 5.23. Since ℓ^L is a Lévy process, for any integers n and L one can decompose $\ell^L(n)$ into

$$\ell^L(n)=\ell^L(0)+\sum_{k=1}^{n}\big\{\ell^L(k)-\ell^L(k-1)\big\},$$

where the n terms in the sum are i.i.d. random variables whose common distribution is that of $\ell^L(1)$ under $\mathbb{P}_0$. Using Bhattacharya's local central limit theorem (see Theorem 1.5 in [**Bha77**]) and the boundedness assumption on $\mathbb{E}_0[|\ell^L(1)|^4]$, we can control the deviation of $p^L(x,n)$ from the corresponding probabilities for Brownian motion up to an error of order $o(n^{-1})$ independent of L. Following Cox and Durrett's arguments, we obtain the desired results for integer times. For arbitrary times t, the Markov property applied to ℓ^L at time $\lfloor t\rfloor$ (plus, for (d), the fact that the variations of ℓ^L are bounded on a time interval $[n,n+1]$) completes the proof. $\square$

Proof of Lemma 5.22. To simplify notation, we shall write $T(d_L)$ instead of $T(d_L,\ell^L)$ in the rest of the proof. For every $L\geq 1$, $x\in\mathbb{T}(L)$ and $\lambda>0$, let us define the following quantities :

$$\begin{aligned}F_L(x,\lambda)&=\mathbb{E}_x\big[\exp(-\lambda T(d_L))\big],\\ G_L(x,\lambda)&=\int_0^\infty e^{-\lambda t}p^L(x,t)dt=\mathbb{E}_x\left[\int_0^\infty e^{-\lambda t}\,\mathbb{I}_{\{\ell^L(t)\in B(0,d_L)\}}dt\right].\end{aligned}$$

Applying the strong Markov property to ℓ^L at time $T(d_L)$ and using a change of variables, we obtain (for any x_L)

$$G_L(x_L,\lambda)=\mathbb{E}_{x_L}\Big[e^{-\lambda T(d_L)}G_L\big(\ell^L(T(d_L)),\lambda\big)\Big]. \tag{5.20}$$

From the results of Lemma 5.23, we can derive the asymptotic behaviour of $G_L(x_L,\lambda)$. To this end, let $(v_L)_{L\geq 1}$ and $(u_L)_{L\geq 1}$ be two sequences growing to infinity such that $v_L(\log L)^{-1/2}\to 0$ and $u_L(\log L)^{-1}\to 0$ as $L\to\infty$. Splitting the integral in the definition of $G_L\big(x_L,\frac{\lambda}{L^2\log L}\big)$ into four pieces, we obtain first by (b) of Lemma 5.23

$$\begin{aligned}
&\frac{1}{d_L^2\log L}\int_{v_L L^2}^{\infty}\exp\left(-\frac{\lambda t}{L^2\log L}\right)p^L(x_L,t)dt\\
&\qquad = \frac{1}{d_L^2\log L}\int_{v_L L^2}^{\infty}\exp\left(-\frac{\lambda t}{L^2\log L}\right)\frac{\pi d_L^2}{L^2}\,(1+\delta_{L,1})\,dt\\
&\qquad = \frac{\pi}{\lambda}\,\exp\left(-\frac{\lambda v_L}{\log L}\right)(1+\delta_{L,1}) = \frac{\pi}{\lambda}\,(1+\delta'_{L,1})
\end{aligned}$$

as $L\to\infty$, where $\delta_{L,1},\delta'_{L,1}\to 0$ uniformly in $x\in\mathbb{T}(L)$. By (a) of Lemma 5.23, we have

$$\begin{aligned}
\frac{1}{d_L^2\log L}\int_{\varepsilon_L L^2}^{v_L L^2}\exp\left(-\frac{\lambda t}{L^2\log L}\right)p^L(x_L,t)\,dt &\le \frac{1}{d_L^2\log L}\frac{C_1 d_L^2}{\lfloor L^2\varepsilon_L\rfloor}\,v_L L^2\\
&\sim \frac{C_1 v_L}{\sqrt{\log L}}\to 0,\quad \text{as } L\to\infty
\end{aligned}$$

by our assumption on v_L. By (c) of Lemma 5.23,

$$\begin{aligned}
&\frac{1}{d_L^2\log L}\quad\int_{u_L(1+|x_L|^2\vee d_L^2)}^{\varepsilon_L L^2}\exp\left(-\frac{\lambda t}{L^2\log L}\right)p^L(x_L,t)\,dt\\
&= \frac{1}{d_L^2\log L}\int_{u_L(1+|x_L|^2\vee d_L^2)}^{\varepsilon_L L^2}\frac{d_L^2}{2\sigma_L^2 t}\,(1+\delta_{L,2})\,dt\\
&= \frac{1}{2\sigma_L^2\log L}\,\Big(2\log L-\log(1+|x_L|^2\vee d_L^2)+\log\varepsilon_L-\log u_L\Big)(1+\delta_{L,2})\\
&= \frac{1-\beta\vee\gamma}{\sigma^2}\,(1+\delta'_{L,2}),
\end{aligned}$$

whenever $\frac{\log^+|x_L|}{\log L}\to\beta$ as L grows to infinity. Here again, $\delta_{L,2},\delta'_{L,2}\to 0$ uniformly in $x\in\mathbb{T}(L)$ as $L\to\infty$. Finally, by (d) of Lemma 5.23, we can write

$$\begin{aligned}
&\frac{1}{d_L^2\log L}\int_0^{u_L(1+|x_L|^2\vee d_L^2)}\exp\left(-\frac{\lambda t}{L^2\log L}\right)\,p^L(x_L,t)dt\\
&\qquad\qquad \le \frac{C_2}{d_L^2\log L}\frac{u_L(1+|x_L|^2\vee d_L^2)}{1+d_L^{-2}|x_L|^2}\to 0,
\end{aligned}$$

independently of $(x_L)_{L\ge 1}$ since d_L does not vanish and $u_L(\log L)^{-1}\to 0$.

Combining the above, we obtain that if $\frac{\log^+|x_L|}{\log L}\to\beta$, then

$$\frac{1}{d_L^2\log L}\,G_L\Big(x_L,\frac{\lambda}{L^2\log L}\Big) = \frac{\pi}{\lambda}+\frac{1-(\beta\vee\gamma)}{\sigma^2}+o(1),\qquad \text{as } L\to\infty,$$

where the remainder does not depend on $(x_L)_{L\ge 1}$. Coming back to (5.20) with $x_L\in\Gamma(L,1)$, the uniform convergence obtained above, together with the fact that $\ell^L(d_L)\in B(0,d_L)$ a.s. yield

$$\lim_{L\to\infty}\mathbb{E}_{x_L}\left[\exp\left(-\frac{\lambda\pi\sigma^2\,T(d_L)}{(1-\gamma)L^2\log L}\right)\right] = \frac{(1-\gamma)/(\sigma^2\lambda)}{(1-\gamma)/(\sigma^2\lambda)+(1-\gamma)/\sigma^2} = \frac{1}{1+\lambda}, \tag{5.21}$$

which we recognize as the Laplace transform of an Exp(1)-random variable. Since the left-hand side of (5.21) is monotone in λ and the function $\lambda\mapsto(1+\lambda)^{-1}$ is continuous, this convergence is in fact uniform in $\lambda\ge 0$. By standard approximation arguments (see for instance the proof of

Theorem 4 in [**Cox89**]), we obtain that for any fixed $t > 0$,

$$\lim_{L\to\infty} \sup_{x_L\in\Gamma(L,1)} \left|\mathbb{P}_{x_L}\left[\frac{\lambda\pi\sigma^2}{(1-\gamma)L^2\log L}\, T(d_L) > t\right] - e^{-t}\right| = 0,$$

and, by monotonicity and the fact that all the quantities involved tend to 0 as $t \to \infty$, this convergence is uniform in $t \geq 0$. The interested reader will find all the missing details in [**CD02**]. □

Another consequence of Lemma 5.23 is the following result, which bounds the probability that ℓ^L hits a ball of bounded radius during a 'short' interval of time.

LEMMA 5.24. *Fix $R > 0$. Let $(U_L)_{L\geq 1}$ and $(u_L)_{L\geq 1}$ be two sequences increasing to infinity such that $U_L L^{-2} \to \infty$ as $L \to \infty$ and $2u_L \leq L^2(\log L)^{-1/2}$ for every $L \geq 1$. Then, there exist $C > 0$ and $L_0 \in \mathbb{N}$ such that for each $L \geq L_0$ and all $x \in \mathbb{T}(L)$,*

$$\mathbb{P}_x\Big[T(R,\ell^L) \in [U_L - u_L, U_L]\Big] \leq \frac{Cu_L}{L^2}.$$

Proof. Let $x \in \mathbb{T}(L)$. Using the strong Markov property at time $T(R,\ell^L)$, we can write

$$\begin{aligned} &\mathbb{P}_x\big[\ell^L(U_L + u_L) \in B(0,R)\big] \\ &\geq \int_{U_L-u_L}^{U_L}\int_{B(0,R)} \mathbb{P}_x\big[T(R,\ell^L)\in ds, \ell^L(s)\in dy\big]\mathbb{P}_y\big[\ell^L(U_L+u_L-s)\in B(0,R)\big]. \end{aligned} \tag{5.22}$$

Note that, on the right-hand side of (5.22), the quantity $U_L + u_L - s$ lies in $[u_L, 2u_L]$. We assumed that $u_L \leq L^2(\log L)^{-1/2}$, and so we can use (c) of Lemma 5.23 with $d_L \equiv R$ and write

$$\lim_{L\to\infty} \sup_{y\in B(0,R)} \sup_{u_L\leq t\leq 2u_L} \left|\frac{2\sigma_L^2 t}{R^2}\,\mathbb{P}_y\big[\ell^L(t)\in B(0,R)\big] - 1\right| = 0,$$

which gives us the existence of a constant C_0 and of an index L_0 such that for each $L \geq L_0$, $y \in B(0,R)$ and $t \in [u_L, 2u_L]$,

$$\mathbb{P}_y\big[\ell^L(t)\in B(0,R)\big] \geq \frac{C_0}{t} \geq \frac{C_0}{2u_L}.$$

Furthermore, since $U_L L^{-2} \to \infty$, we can use (b) of Lemma 5.23 to obtain the existence of $C_1 > 0$ and $L_1 \in \mathbb{N}$ such that for every $L \geq L_1$ and all $y \in \mathbb{T}(L)$,

$$\mathbb{P}_y\big[\ell^L(U_L + u_L)\in B(0,R)\big] \leq \frac{C_1}{L^2}.$$

Using these two inequalities in (5.22), we have for L large enough and for all $x \in \mathbb{T}(L)$

$$\frac{C_1}{L^2} \geq \mathbb{P}_x\big[T(R,\ell^L)\in[U_L-u_L,U_L]\big] \times \frac{C_0}{2u_L},$$

which gives us the desired result. □

6. Proof of Theorem 5.11

Armed with the estimates of Section 5, we can now turn to the proofs of our main results.

NOTATION 5.25. *For each $L \geq 1$, let $\{\xi^L(t), t \geq 0\}$ be the Lévy process on $\mathbb{T}(L)$ whose distribution is the same as that of the motion of a single lineage subject to the large and small reproduction events generated by Π_L^s and Π_L^B.*

In the rest of this section, we assume that the assumptions of Theorem 5.11 are satisfied.

6.1. Coalescence time for two lineages. We begin by studying the genealogical process of a pair of lineages starting at distance $\mathcal{O}(L)$ from each other. Since the motions ξ_1^L and ξ_2^L of the lineages are distributed like two independent copies of the process ξ^L until the random time T_L at which they come at distance less than $2R^B\psi(L)$, the difference

$$X^L(t) := \xi_1^L(t) - \xi_2^L(t), \qquad 0 \le t \le T_L$$

has the same distribution as $\big\{\xi^L(2t),\ 0 \le t \le \frac{1}{2}\, T(2R^B\psi(L), \xi^L)\big\}$. We shall use Lemma 5.22 to derive the limiting distribution of T_L, but we first need to introduce the relevant variances. Consider a single lineage. Because it jumps at a finite rate owing to small and large events, the two following quantities are well-defined and finite :

$$\text{(5.23)} \qquad \sigma_s^2 := \int y^2\, \chi^s(dy, dz) \qquad \text{and} \qquad \sigma_B^2 := \int y^2\, \chi^B(dy, dz),$$

where χ^s stands for the intensity measure of the small jumps experienced by the lineage and χ^B for that of the large jumps rescaled by $\psi(L)^{-1}$ (the form of these two measures is given in (5.7)). We now have all the ingredients we need to describe the asymptotic 'gathering time' of two lineages.

PROPOSITION 5.26. *(a) If $\rho(L)\psi(L)^{-2} \to 0$ as $L \to \infty$, then*

$$\lim_{L\to\infty} \sup_{t\ge 0} \sup_{A_L \in \Gamma_{\mathcal{A}}(L,2)^*} \left| \mathbb{P}_{A_L}\left[T_L > \frac{(1-\alpha)\rho(L) L^2 \log L}{2\pi \sigma_B^2 \psi(L)^2}\, t \right] - e^{-t} \right| = 0.$$

(b) If $\rho(L)^{-1}\psi(L)^2 \to b \in [0,\infty)$ as $L \to \infty$, then

$$\lim_{L\to\infty} \sup_{t\ge 0} \sup_{A_L \in \Gamma_{\mathcal{A}}(L,2)^*} \left| \mathbb{P}_{A_L}\left[T_L > \frac{(1-\alpha) L^2 \log L}{2\pi (\sigma_s^2 + b\sigma_B^2)}\, t \right] - e^{-t} \right| = 0.$$

Proof. Let us first recall two results on Poisson point processes, which are consequences of the exponential formula given, for instance, in Section 0.5 of [**Ber96**]. Following Bertoin's notation, let $\{e(t), t \ge 0\}$ be a Poisson point process on $\mathbb{R} \times \mathbb{R}_+$ with intensity measure $\kappa(dy) \otimes dt$, where the Borel measure κ satisfies

$$\text{(5.24)} \qquad \int_{\mathbb{R}} |1 - e^y| \kappa(dy) < \infty \qquad \text{and} \qquad \int_{\mathbb{R}} y^m \kappa(dy) = 0, \quad m \in \{1, 3\}.$$

Under these conditions, we have for each fixed $t > 0$

$$\text{(5.25)} \qquad \mathbb{E}\left[\Big(\sum_{s\le t} e(s)\Big)^2\right] = t \int_{\mathbb{R}} y^2 \kappa(dy),$$

$$\text{(5.26)} \qquad \mathbb{E}\left[\Big(\sum_{s\le t} e(s)\Big)^4\right] = 3t^2 \Big(\int_{\mathbb{R}} y^2 \kappa(dy)\Big)^2 + t \int_{\mathbb{R}} y^4 \kappa(dy).$$

These properties will be useful in computing the variances and fourth moments of the random variables considered below.

Let us start with the proof of (a). Consider the process ℓ^L defined by: For every $t \ge 0$,

$$\ell^L(t) = \frac{1}{\psi(L)}\, \xi^L\big(2\rho(L)t\big).$$

This process evolves on the torus of size $\psi(L)^{-1}L$, and makes jumps of size $\mathcal{O}(\psi(L)^{-1})$ at a rate of order $\mathcal{O}(\rho(L))$, as well as jumps of size $\mathcal{O}(1)$ at a rate of order $\mathcal{O}(1)$.

Let us check that ℓ^L satisfies the assumptions of Lemma 5.22. To this end, we view $\ell^L(1)$ starting at 0 as the sum of its jumps and adapt the problem to use the results on Poisson point processes given above. First, let us define $\hat{\ell}^L$ as the Lévy process on $\mathbb{R}^2$ evolving like ℓ^L (but without periodic conditions). For $i \in \{0, 1\}$ and each $L \ge 1,\ t \ge 0$, let $\hat{\ell}^{L,i}(t)$ denote the i-th coordinate of $\hat{\ell}^L(t)$. Note that the distance reached by ℓ^L up to a given time t is less than or equal to the distance at which $\hat{\ell}^L$ traveled up to t, and so we can write

$$\begin{aligned}\mathbb{E}_0\big[|\ell^L(1)|^4\big] \leq \mathbb{E}_0\big[|\hat{\ell}^L(1)|^4\big] &= \mathbb{E}_0\Big[\Big\{\hat{\ell}^{L,1}(1)^2+\hat{\ell}^{L,2}(1)^2\Big\}^2\Big]\\ &\leq 2\Big\{\mathbb{E}_0\big[\hat{\ell}^{L,1}(1)^4\big]+\mathbb{E}_0\big[\hat{\ell}^{L,2}(1)^4\big]\Big\}.\end{aligned}$$

By symmetry, we need only bound $\mathbb{E}_0\big[\hat{\ell}^{L,1}(1)^4\big]$. Denote by $a_1, a_2, \ldots \in [-2R^s, 2R^s]^2$ (resp., $b_1, b_2, \ldots \in [-2R^B\psi(L), 2R^B\psi(L)]^2$) the sequence of the jumps of $\hat{\ell}^{L,1}$ before time 1 due to small (resp., large) events. Using the convexity of $y \mapsto y^4$, we have

$$\mathbb{E}_0\big[\hat{\ell}^{L,1}(1)^4\big] = \mathbb{E}_0\Big[\Big(\sum_i a_i + \sum_j b_j\Big)^4\Big] \leq 8\,\mathbb{E}_0\Big[\Big(\sum_i a_i\Big)^4 + \Big(\sum_j b_j\Big)^4\Big]. \tag{5.27}$$

Applying (5.26) to each term on the right-hand side of (5.27) yields

$$\mathbb{E}_0\big[(\hat{\ell}^{L,1}(1))^4\big] \leq 96\frac{\rho(L)^2}{\psi(L)^4}\,\sigma_s^4 + 16\frac{\rho(L)}{\psi(L)^4}\int y^4\chi^s(dy,dz) + 96\sigma_B^4 + 16\int y^4\chi^B(dy,dz), \tag{5.28}$$

which is bounded uniformly in L since $\rho(L)\psi(L)^{-2}$ vanishes as L grows to infinity, and each integral is finite. Coming back to the original problem, we obtain that Assumption 5.21 (ii) holds for the sequence of processes $(\ell^L)_{L\geq 1}$.

As concerns Assumption 5.21 (i), observe that σ_L^2 is simply the variance of $\ell^{L,1}(1)$. To obtain the asymptotic behaviour of σ_L^2, we show that up to time 1, ℓ^L does not see that it is on a torus. Hence, with high probability $\ell^{L,1}(1)^2 = \hat{\ell}^{L,1}(1)^2$ and so

$$\mathbb{E}_0\big[\ell^{L,1}(1)^2\big] \approx \mathbb{E}_0\big[\hat{\ell}^{L,1}(1)^2\big] = 2\frac{\rho(L)}{\psi(L)^2}\int y^2\chi^s(dy,dz) + 2\int y^2\chi^B(dy,dz) = 2\sigma_B^2 + o(1)$$

as $L \to \infty$, where the second equality uses (5.25). To make the first equality rigorous, we apply Doob's maximal inequality to the submartingale $|\hat{\ell}^L|^4$. This yields, with a constant $C > 0$ which may change from line to line,

$$\mathbb{P}_0\Big[\sup_{0\leq s\leq 1}|\hat{\ell}^L(s)| > \frac{L}{3\psi(L)}\Big] \leq \frac{C\psi(L)^4}{L^4}\,\mathbb{E}_0\big[|\hat{\ell}^L(1)|^4\big].$$

But the calculation in (5.28) shows that the latter expectation is finite, and so

$$\mathbb{P}_0\Big[\sup_{0\leq s\leq 1}|\hat{\ell}^L(s)| > \frac{L}{3\psi(L)}\Big] \leq C\frac{\psi(L)^4}{L^4}. \tag{5.29}$$

On the event $\mathcal{E}_L := \big\{\sup_{0\leq s\leq 1}|\hat{\ell}^L(s)| \leq \frac{L}{3\psi(L)}\big\}$, the paths of ℓ^L and $\hat{\ell}^L$ can be coupled so that $\ell^L(s) = \hat{\ell}^L(s)$ for every $s \in [0,1]$, and since these quantities are bounded for each L we can write

$$\begin{aligned}\mathbb{E}_0\big[(\ell^{L,1}(1))^2\big] &= \mathbb{E}_0\big[(\hat{\ell}^{L,1}(1))^2\,\mathbb{I}_{\mathcal{E}_L}\big] + \mathbb{E}_0\big[(\ell^{L,1}(1))^2\,\mathbb{I}_{\mathcal{E}_L^c}\big]\\ &= \mathbb{E}_0\big[(\hat{\ell}^{L,1}(1))^2\big] - \mathbb{E}_0\big[(\hat{\ell}^{L,1}(1))^2\,\mathbb{I}_{\mathcal{E}_L^c}\big] + \mathbb{E}_0\big[(\ell^{L,1}(1))^2\,\mathbb{I}_{\mathcal{E}_L^c}\big].\end{aligned} \tag{5.30}$$

By (5.29) and the fact that ℓ^L evolves on the torus of size $L\psi(L)^{-1}$, the last term on the right-hand side of (5.30) is bounded by

$$C\,\frac{L^2}{\psi(L)^2} \times \frac{\psi(L)^4}{L^4} = C\,\frac{\psi(L)^2}{L^2} \to 0 \qquad \text{as } L \to \infty.$$

For the second term on the right-hand side of (5.30), let $\hat{s}_L(1) := \sup_{0\leq s\leq 1}|\hat{\ell}^L(s)|$. Using Fubini's theorem on the second line, we have

$$\begin{aligned}\mathbb{E}_0\big[(\hat{\ell}^{L,1}(1))^2\,\mathbb{I}_{\mathcal{E}_L^c}\big] &\leq \mathbb{E}_0\big[\hat{s}_L(1)^2\,\mathbb{I}_{\mathcal{E}_L^c}\big]\\ &= \int_0^\infty \mathbb{P}_0\Big[\hat{s}_L(1) > \frac{L}{3\psi(L)}\vee\sqrt{y}\Big]\,dy\\ &= \frac{L^2}{9\psi(L)^2}\,\mathbb{P}_0\Big[\hat{s}_L(1) > \frac{L}{3\psi(L)}\Big] + \int_{\frac{L^2}{9\psi(L)^2}}^\infty \mathbb{P}_0\big[\hat{s}_L(1) > \sqrt{y}\big]\,dy. \qquad (5.31)\end{aligned}$$

Now, by the argument leading to (5.29), $\mathbb{P}_0[\hat{s}_L(1) > \sqrt{y}]$ is bounded by Cy^{-2} for each $y > 0$, where C is a constant independent of y. Consequently, the right-hand side of (5.31) is bounded by

$$C'\,\frac{\psi(L)^2}{L^2} + C\int_{L^2/(9\psi(L)^2)}^\infty \frac{dy}{y^2}\ \to\ 0 \qquad \text{as } L\to\infty.$$

Coming back to (5.30), we can conclude than

$$\sigma_L^2 = 2\sigma_B^2 + o(1) \qquad \text{as } L\to\infty.$$

If we now recall the equality in distribution described at the beginning of the section, we can use Lemma 5.22 applied to ℓ^L on the torus of size $L\psi(L)^{-1}$ and the entrance time into $B(0,2R^B)$ to write that

$$\lim_{L\to\infty}\sup_{t\geq 0}\sup_{A_L\in\Gamma_{\mathcal{A}}(L,2)^*}\left|\mathbb{P}_{A_L}\left[T_L > \frac{L^2\psi(L)^{-2}\log(L\psi(L)^{-1})}{2\pi\sigma_B^2}\,t\right] - e^{-t}\right| = 0. \qquad (5.32)$$

By our assumption on $|\alpha\log L - \log(L\psi(L)^{-1})|$ and Lemma 5.24 applied to ℓ^L, (a) of Proposition 5.26 follows from (5.32).

Let us now turn to the proof of (b). This time, we define ℓ^L for every $t\geq 0$ by

$$\ell^L(t) = \frac{1}{\psi(L)}\,\xi^L(2\psi(L)^2 t).$$

Similar calculations give, as $L\to\infty$,

$$E_0\left[|\ell^L(1)|^2\right] = 2\sigma_s^2 + 2b\sigma_B^2 + o(1) \qquad \text{if } \rho(L)^{-1}\psi(L)^2 \to b\in[0,\infty).$$

and $E_0\left[|\ell^L(1)|^4\right]$ is bounded uniformly in L. We can therefore apply Lemma 5.22 to ℓ^L as above. ☐

Having established the time that it takes for two lineages starting from distance L apart to come close enough together that they have a chance to coalesce, we now calculate the additional time required for them to actually do so. We shall have to distinguish between several regimes, depending on whether large or small events prevail in the evolution of the pair of lineages. Our goal in the rest of this section is to prove the following result.

THEOREM 5.27. *For each $L\in\mathbb{N}$, let t_L denote the coalescence time of the pair of lineages under consideration. Then,*

(a) If $\frac{\psi(L)^2}{\rho(L)}\to\infty$ as $L\to\infty$,

$$\lim_{L\to\infty}\sup_{t\geq 0}\sup_{A_L\in\Gamma_{\mathcal{A}}(L,2)^*}\left|\mathbb{P}_{A_L}\left[t_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\,t\right] - e^{-t}\right| = 0.$$

(b) If $\frac{\psi(L)^2}{\rho(L)}\to b\in[0,\infty)$ and $\frac{\psi(L)^2\log L}{\rho(L)}\to\infty$ as $L\to\infty$,

$$\lim_{L\to\infty}\sup_{t\geq 0}\sup_{A_L\in\Gamma_{\mathcal{A}}(L,2)^*}\left|\mathbb{P}_{A_L}\left[t_L > \frac{(1-\alpha)L^2\log L}{2\pi(\sigma_s^2+b\sigma_B^2)}\,t\right] - e^{-t}\right| = 0.$$

(c) *If* $\frac{\psi(L)^4}{\rho(L)} \rightarrow \mathfrak{b} \in [0,\infty)$ *or* $\frac{L^2 \log L}{\rho(L)} \rightarrow 0$ *as* $L \rightarrow \infty$ *(and so* $\frac{\psi(L)^2 \log L}{\rho(L)} \rightarrow 0$*), then*

$$\lim_{L\rightarrow\infty} \sup_{t\geq 0} \sup_{A_L\in\Gamma_{\mathcal{A}}(L,2)^*} \left|\mathbb{P}_{A_L}\left[t_L > \frac{L^2\log L}{2\pi\sigma_s^2}\, t\right] - e^{-t}\right| = 0.$$

The cases (a) and (b) are separated only because the timescales of interest are not of the same order, but the reasons why they hold are identical: in both cases, large jumps are frequent enough that, once the lineages have been gathered at distance $2R^B\psi(L)$, they coalesce in a time negligible compared to T_L. In contrast, in (c) we assume that the rate at which the lineages are affected by large events is so slow that we have to wait for the lineages to be gathered at distance less than $2R^s$ before they have a chance to coalesce (and they do so in a negligible time compared to $L^2\log L$). If none of the above conditions hold, then the proof of (c) will show that, also in this case, the probability that a large event affects the lineages when they are at distance less than $2R^B\psi(L)$ and before a time of order $\mathcal{O}(L^2\log L)$ vanishes as L tends to infinity. However, we are no longer able to describe precisely the limiting behaviour of t_L, see the remark at the end of the proof of (c).

Let us first make more precise the sense in which the additional time to coalescence is negligible once the lineages have been gathered at the right distance.

PROPOSITION 5.28. (a) *If* $\frac{\psi(L)^2\log L}{\rho(L)} \rightarrow \infty$ *as* $L\rightarrow\infty$*, then*

$$\lim_{L\rightarrow\infty} \sup_{A_L} \mathbb{P}_{A_L}\left[t_L > \frac{L^2}{\log L}\Big(\frac{\rho(L)}{\psi(L)^2}\wedge 1\Big)\right] = 0, \tag{5.33}$$

where the supremum is taken over all samples $A_L = \big\{(\{1\},x_1^L),(\{2\},x_2^L)\big\}$ *such that* $|x_1^L - x_2^L| \leq 2R^B\psi(L)$.

(b) *Let* $(\Phi(L))_{L\geq 1}$ *be a sequence such that* $\Phi(L)\rightarrow\infty$ *as* $L\rightarrow\infty$*. Under any condition, we have*

$$\lim_{L\rightarrow\infty} \sup_{A'_L} \mathbb{P}_{A'_L}\big[t_L > \Phi(L)\big] = 0, \tag{5.34}$$

where the supremum is now taken over all samples $A'_L = \big\{(\{1\},x_1^L),(\{2\},x_2^L)\big\}$ *such that* $|x_1^L - x_2^L| \leq 2R^s$.

The result in (a) shows that, under the condition specified, the coalescence time of two lineages at distance at most $2R^B\psi(L)$ is indeed much smaller than T_L (which is of order $L^2\log L \times \big(1\wedge \rho(L)\psi(L)^{-2}\big)$ by Proposition 5.26).

Proof of Proposition 5.28. Recall that for each $L\in\mathbb{N}$, we defined X^L as the difference between the locations of the lineages ξ_1^L and ξ_2^L on the torus $\mathbb{T}(L)$. In the following, if both lineages are affected by the same event, we shall consider that X^L hits 0 but the number of lineages remains equal to 2, which means that they can separate again later (if the measures ν_r^s and ν_r^B are not all the point mass at 1). However, it is the first time at which such an event occurs which will be of interest, and we keep the notation t_L to denote this time. As we already noticed, X^L behaves like $\{\xi^L(2t),\ t\geq 0\}$ outside $B\big(0,2R^B\psi(L)\big)$, whereas inside the ball it can hit 0 owing to reproduction events affecting both lineages ξ_1^L and ξ_2^L.

Case (a)**.** For each $L\in\mathbb{N}$, set $q_0^L = Q_0^L := 0$ and for every $i\geq 1$,

$$Q_i^L := \inf\Big\{t > q_{i-1}^L :\ X^L(t)\notin B\Big(0,\frac{7}{4}R^B\psi(L)\Big)\Big\}$$

and

$$q_i^L := \inf\Big\{t > Q_i^L :\ X^L(t)\in B\Big(0,\frac{3}{2}R^B\psi(L)\Big)\Big\},$$

with the convention that inf $\emptyset = +\infty$. We shall use the following lemmas, which will enable us to describe how X^L wanders around in $\mathbb{T}(L)$, independently of whether it ever hits 0 or not.

LEMMA 5.29. *There exist a function $g : \mathbb{R}_+ \to \mathbb{R}_+$ vanishing at infinity, $C_q > 0$, $u_q > 1$ and $L_q \in \mathbb{N}$ such that for every $L \geq L_q$ and $u \geq u_q$,*

$$\sup_{x\in B(0,4R^B)\setminus B(0,(7/4)R^B)} \mathbb{P}_{\psi(L)x}\big[q_1^L > \rho(L)u\big] \leq g(u) \qquad \text{if } \rho(L) = \mathcal{O}(\psi(L)^2),$$

$$\sup_{x\in B(0,4R^B)\setminus B(0,(7/4)R^B)} \mathbb{P}_{\psi(L)x}\big[q_1^L > \psi(L)^2u\big] \leq \frac{C_q}{\log u} \qquad \text{if } \rho(L)^{-1}\psi(L)^2 \to 0.$$

LEMMA 5.30. *Suppose that*

$$\mathrm{Leb}\big(\{r \in [0,R^B] :\ \nu_r^B \notin \{\delta_0,\delta_1\}\}\big) > 0. \tag{5.35}$$

Then, there exists a constant $C_Q < \infty$ such that for each $L \geq 1$,

$$\sup_{x\in B(0,(3/2)R^B)} \frac{1}{\rho(L)}\, \mathbb{E}_{\psi(L)x}\big[Q_1^L\big] < C_Q.$$

Condition (5.35) guarantees that, whenever X^L hits 0, it has a chance not to remain stuck at this value for all times.

LEMMA 5.31. *Suppose that $\rho(L)\psi(L)^{-2}$ has a finite limit (which can be 0) as $L \to \infty$. Then, there exists $\theta_1 \in (0,1)$ such that for every $L \geq 1$,*

$$\inf_{x\in B(0,(3/2)R^B)} \mathbb{P}_{\psi(L)x}\big[X^L \text{ hits 0 before leaving } B\big(0,(7/4)R^B\psi(L)\big)\big] \geq \theta_1. \tag{5.36}$$

If $\rho(L)^{-1}\psi(L)^2 \to 0$ as $L \to \infty$, there exist $\theta_2 \in (0,1)$ and $\theta_3 > 0$ such that

$$\begin{aligned}\inf_{x\in B(0,(3/2)R^B)} &\mathbb{P}_{\psi(L)x}\big[X^L \text{ hits 0 before leaving } B\left(0,(7/4)R^B\psi(L)\right)\big]\\ &\geq \theta_2\bigg(1-\exp\Big\{-\theta_3\frac{\psi(L)^2}{\rho(L)}\Big\}\bigg).\end{aligned} \tag{5.37}$$

We postpone the proof of these lemmas until after the proof of Theorem 5.27. The following technique is inspired by that used in [**CD02**] and [**ZCD05**], although the motions of the lineages and the mechanism of coalescence here are more complex and require slightly more work. For the sake of clarity, we show (5.33) in the case where $\rho(L)\psi(L)^{-2}$ has a finite limit, and then comment on how to adapt the arguments when the limit is infinite.

Assume first that Condition (5.35) holds. Recall the definition of Q_i^L and q_i^L given above, and define k_L by

$$k_L := \max\bigg\{n :\ Q_n^L \leq \frac{\rho(L)L^2}{\psi(L)^2\log L}\bigg\}.$$

By Lemma 5.31, there exists $\theta_1 > 0$ such that for every $L \geq 1$ and $x \in B(0,(3/2)R^B\psi(L))$,

$$\mathbb{P}_x\big[X^L \text{ hits 0 before leaving } B\big(0,(7/4)R^B\psi(L)\big)\big] \geq \theta_1.$$

Hence, for every $x \in B\big(0,2R^B\psi(L)\big)$, we have

$$\mathbb{P}_x\Big[t_L > \frac{\rho(L)L^2}{\psi(L)^2\log L}\Big] \leq \mathbb{P}_x\big[t_L > Q_{k_L}^L\big] \leq \mathbb{E}_x\big[(1-\theta_1)^{k_L}\big]. \tag{5.38}$$

Let us fix $x \in B\big(0,2R^B\psi(L)\big)$ and show that $k_L \to \infty$ as $L \to \infty$, in $\mathbb{P}_x$-probability. The fact that the bounds obtained below do not depend on $x \in B\big(0,2R^B\psi(L)\big)$ will then give us the desired uniformity. Let $M \in \mathbb{N}$. We have

$$\mathbb{P}_x\big[k_L < M\big] = \mathbb{P}_x\Big[Q_M^L > \frac{\rho(L)L^2}{\psi(L)^2\log L}\Big]$$

$$\begin{aligned}
&= \mathbb{P}_x\Big[\sum_{i=1}^{M}(Q_i^L - q_{i-1}^L) + \sum_{i=1}^{M-1}(q_i^L - Q_i^L) > \frac{\rho(L)L^2}{\psi(L)^2\log L}\Big] \\
&\leq \sum_{i=1}^{M}\mathbb{P}_x\Big[Q_i^L - q_{i-1}^L > \frac{\rho(L)L^2}{2M\psi(L)^2\log L}\Big] \\
&\quad + \sum_{i=1}^{M-1}\mathbb{P}_x\Big[q_i^L - Q_i^L > \frac{\rho(L)L^2}{2(M-1)\psi(L)^2\log L}\Big], \qquad (5.39)
\end{aligned}$$

where the last inequality uses the fact that at least one of the $2M-1$ terms of the sums on the second line must be larger than a fraction $(2M-1)^{-1}$ of the total time. Now, using the Markov inequality, the strong Markov property at time q_{i-1}^L and then Lemma 5.30, we can write for each i

$$\begin{aligned}
\mathbb{P}_x\Big[Q_i^L - q_{i-1}^L > \frac{\rho(L)L^2}{2M\psi(L)^2\log L}\Big] &\leq \frac{2M\psi(L)^2\log L}{\rho(L)L^2}\,\mathbb{E}_x\big[Q_i^L - q_{i-1}^L\big] \\
&\leq \frac{2M\psi(L)^2\log L}{\rho(L)L^2}\sup_{y\in B(0,(3/2)R^B)}\mathbb{E}_{\psi(L)y}[Q_1^L] \\
&\leq \frac{2C_Q M\psi(L)^2\log L}{L^2}.
\end{aligned}$$

If we now apply the strong Markov property to X^L at time Q_i^L and use Lemma 5.29 together with the fact that $X^L(Q_i^L) \in B(0,4R^B\psi(L))$ with probability one, we obtain for each i, and L large enough

$$\mathbb{P}_x\Big[q_i^L - Q_i^L > \frac{\rho(L)L^2}{2(M-1)\psi(L)^2\log L}\Big] \leq g\Big(\frac{L^2}{2(M-1)\psi(L)^2\log L}\Big).$$

Coming back to (5.39), we obtain

$$\begin{aligned}
\mathbb{P}_x\big[k_L < M\big] &\leq \frac{2M^2C_Q\psi(L)^2\log L}{L^2} + Mg\Big(\frac{L^2}{2(M-1)\psi(L)^2\log L}\Big) \\
&\to 0, \qquad \text{as } L\to\infty.
\end{aligned}$$

To complete the proof of (a) when Condition (5.35) holds and $\rho(L)\psi(L)^{-2}$ has a finite limit, let $\varepsilon > 0$ and fix $M = M(\varepsilon) \in \mathbb{N}$ such that

$$(1-\theta_1)^M < \varepsilon.$$

Splitting the expectation in (5.38) into the integral over $\{k_L \geq M\}$ and $\{k_L < M\}$ yields

$$\limsup_{L\to\infty}\sup_{x\in B(0,2R^B\psi(L))}\mathbb{P}_x\Big[t_L > \frac{\rho(L)L^2}{\psi(L)^2\log L}\Big] \leq \varepsilon + \limsup_{L\to\infty}\sup_{x\in B(0,2R^B\psi(L))}\mathbb{P}_x\big[k_L < M\big] = \varepsilon,$$

and since ε was arbitrary, the desired result follows.

When Condition (5.35) is fulfilled but $\rho(L)\psi(L)^{-2}$ has an infinite limit (and $\frac{\psi(L)^2\log L}{\rho(L)} \to \infty$) as $L\to\infty$, we can apply the same technique to obtain (5.33). This time, we define k_L by

$$k_L := \max\Big\{n:\ Q_n^L \leq \frac{L^2}{\log L}\Big\},$$

and use Lemma 5.31 to write as in (5.38) that, for every $x \in B\big(0,2R^B\psi(L)\big)$,

$$\mathbb{P}_x\Big[t_L > \frac{L^2}{\log L}\Big] \leq \mathbb{E}_x\Big[\Big(1-\theta_2\Big(1-\exp\Big\{-\theta_3\frac{\psi(L)^2}{\rho(L)}\Big\}\Big)\Big)^{k_L}\Big].$$

If $(\varepsilon_L)_{L\geq 1}$ is such that $\varepsilon_L \to 0$ but $\frac{\varepsilon_L \psi(L)^2 \log L}{\rho(L)} \to \infty$ as $L\to\infty$, the same arguments as above (using the second part of Lemma 5.29) yield, for L large enough,

$$\sup_{x\in B(0,2R^B\psi(L))} \mathbb{P}_x\big[k_L < \varepsilon_L \log L\big] \leq \frac{2C_Q\varepsilon_L^2\rho(L)(\log L)^3}{L^2} \\ +(\varepsilon_L \log L)C_q\bigg\{\log\bigg(\frac{L^2}{2\psi(L)^2\varepsilon_L(\log L)^2}\bigg)\bigg\}^{-1},$$

which tends to 0 as L tends to infinity since the term within the curly brackets is of order $\mathcal{O}(\log L)$ when L is large and the first time is negligible compared to $L^{-2}\psi(L)^2(\log L)^4$ by our assumption $\rho(L) \ll \psi(L)^2\log L$. Since by our choice of $(\varepsilon_L)_{L\geq 1}$

$$\log\bigg\{\Big(1-\theta_2\Big(1-\exp\Big\{-\theta_3\frac{\psi(L)^2}{\rho(L)}\Big\}\Big)\Big)^{\varepsilon_L\log L}\bigg\} = -\theta_2\theta_3(\varepsilon_L\log L)\frac{\psi(L)^2}{\rho(L)}\,(1+o(1)) \to -\infty,$$

for each $\varepsilon>0$ we can find $L_0(\varepsilon)\in\mathbb{N}$ such that for all $L\geq L_0(\varepsilon)$,

$$\Big(1-\theta_2\Big(1-\exp\Big\{-\theta_3\frac{\psi(L)^2}{\rho(L)}\Big\}\Big)\Big)^{\varepsilon_L\log L} \leq \varepsilon,$$

and conclude in the same manner as before.

Let us finish the proof of (a) by removing the Assumption (5.35). In the preceding proof, the main idea is that each time X^L passes through $B\big(0,(3/2)R^B\psi(L)\big)$, the two lineages have an opportunity to try to coalesce and their success probability is bounded from below by the quantity obtained in Lemma 5.31. However, if we do not assume that (5.35) holds, X^L may become stuck at 0 once it has hit it, and so the number k_L of such sojourns in $B\big(0,(3/2)R^B\psi(L)\big)$ may be finite. This makes our arguments break down. Nevertheless, X^L can only hit 0 through a coalescence event, and so this issue is merely an artefact of the technique of the proof. To overcome it, let us increase the rate of reproduction events by a factor 2, but divide each probability to be affected by 2. Overall, coalescence will take a longer time in this new setting, but the motions of the lineages before their coalescence time will remain identical in distribution.

More precisely, assume that (5.35) does not hold. Define $\hat{\Pi}_L^B$ as a Poisson point process on $\mathbb{R}\times\psi(L)^{-1}\mathbb{T}(L)\times(0,\infty)$, independent of Π_L^s and Π_L^B and with intensity measure $2dt\otimes dx\otimes\mu^B(dr)$, and for each $r>0$ such that $\nu_r^B=\delta_1$, set $\hat{\nu}_r^B := \delta_{1/2}$. Let also $\hat{\Pi}_L^s$ be a Poisson point process with the same distribution as Π_L^s and independent of all the other point processes. Call $\hat{X}^L$ the process defined in the same manner as X^L but with Π_L^B (resp., Π_L^s, $\nu_r^B:=\delta_1$) replaced by $\hat{\Pi}_L^B$ (resp., $\hat{\Pi}_L^s$, $\hat{\nu}_r^B:=\delta_{1/2}$). By computing the intensity of the jumps of a single lineage, one can observe that it is equal to

$$dt\otimes\bigg(\frac{2}{\rho(L)}\int_{|x|/2}^{R^B}\frac{L_r(x)}{\pi r^2}\times\frac{1}{2}\,\mu^B(dr)d(\psi(L)x)+\int_{|x|/2}^{R^s}\int_0^1\frac{L_r(x)}{\pi r^2}\,u\,\nu_r^s(du)\mu^s(dr)dx\bigg),$$

which is precisely that of ξ^L. If we now compute the coalescence rate of two lineages at distance $z\in[0,2R^B\psi(L)]$, we obtain the same term due to small events for X^L and $\hat{X}^L$, to which is added the contribution of large events, i.e.

$$\frac{1}{\rho(L)}\int_{z/2}^{R^B}L_r(z)\mu^B(dr)\quad\text{for } X^L,\qquad\text{and}\qquad\frac{1}{4\rho(L)}\int_{z/2}^{R^B}L_r(z)\mu^B(dr)\quad\text{for } \hat{X}^L.$$

Hence, the evolutions of both processes follow the same law outside $B(0,2R^B\psi(L))$, the contribution of large events whose area encompasses only one of the two lineages is identical even within $B(0,2R^B\psi(L))$, and coalescence occurs at a higher rate for X^L than for $\hat{X}^L$. This gives us for every $L \geq 1$ and $x \in \mathbb{T}(L)$,

$$\mathbb{P}_x\Big[t_L > \frac{L^2}{\log L}\Big(\frac{\rho(L)}{\psi(L)^2} \wedge 1\Big)\Big] \leq \mathbb{P}_x\Big[\hat{t}_L > \frac{L^2}{\log L}\Big(\frac{\rho(L)}{\psi(L)^2} \wedge 1\Big)\Big],$$

where $\hat{t}_L$ is defined in an obvious manner. But Condition (5.35) holds for $\hat{X}^L$, and so we can use the result obtained in the previous paragraph to complete the proof of (a) when (5.35) does not hold.

Case (b)**.** The arguments are essentially the same. First of all, since we assumed that $\rho(L)$ grows to infinity as $L \rightarrow \infty$, and because

$$\mathbb{P}_x\big[t_L > \Phi(L)\big] \leq \mathbb{P}_x\big[t_L > \Phi'(L)\big]$$

whenever $\Phi(L) \geq \Phi'(L)$, we can restrict our attention to sequences $(\Phi(L))_{L\geq 1}$ s.t. $\Phi(L)/\rho(L) \rightarrow 0$ as $L \rightarrow \infty$. Let $\mathcal{E}_L$ denote the event that no large events affected any of the lineages before time $\Phi(L)$. Let $\theta_{\max} \in (0,\infty)$ be such that the maximal rate at which at least one of the two lineages of the sample is affected by a large event is less than $\theta_{\max}\rho(L)^{-1}$ (recall that the total rate at which at least one of two lineages is affected is smaller than twice the corresponding rate for a single lineage, which is finite and independent of the location of the lineage). For each $L \in \mathbb{N}$, define e_L as an exponential random variable, with parameter $\theta_{\max}\rho(L)^{-1}$. By our assumption on $\Phi(L)$, we can write

$$\mathbb{P}_x[\mathcal{E}_L^c] \leq \mathbb{P}[e_L \leq \Phi(L)] = 1 - \exp\bigg\{-\frac{\theta_{\max}\Phi(L)}{\rho(L)}\bigg\} \rightarrow 0, \quad \text{as } L \rightarrow \infty.$$

The distribution of the process X^L up to the first time at which it is affected by a large event is equal to that of the process $\tilde{X}^L$ (experiencing only small events), up to the random time $e(\tilde{X}^L)$, and so if $\rho(L)^{-1}\theta_{B,L}(x)$ is the rate at which at least one of two lineages at separation $x \in \mathbb{T}(L)$ is affected by a large event, then for each $t \geq 0$ and $y \in \mathbb{T}(L)$

$$\mathbb{P}_y\big[e(\tilde{X}^L) > t\big] = \mathbb{E}_y\bigg[\exp\bigg\{-\int_0^t \frac{\theta_{B,L}\big(\tilde{X}^L(s)\big)}{\rho(L)}\,ds\bigg\}\bigg].$$

By the definition of $\theta_{\max}$, for each $L \in \mathbb{N}$ the variable e_L is stochastically bounded by $e(\tilde{X}^L)$. Consequently, we have for each $x \in B(0,2R^s)$

$$\begin{aligned}\mathbb{P}_x\big[t_L \geq \Phi(L)\big] &\leq \mathbb{P}_x\big[t_L \geq \Phi(L);\ \mathcal{E}_L\big] + \mathbb{P}_x\big[\mathcal{E}_L^c\big]\\ &= \mathbb{P}_x\big[\tilde{t}_L \geq \Phi(L);\ e(\tilde{X}^L) > \Phi(L)\big] + o(1)\\ &\leq \mathbb{P}_x\big[\tilde{t}_L \geq \Phi(L);\ e_L > \Phi(L)\big] + o(1)\\ &= \mathbb{P}_x\big[\tilde{t}_L \geq \Phi(L)\big] + o(1) \quad \text{as } L \rightarrow \infty,\end{aligned}$$

where the remaining terms converge to 0 uniformly in $x \in \mathbb{T}(L)$. Then, an easy modification of the proof of (a) with '$\psi(L) = \rho(L) = 1$' yields the desired result and completes the proof of Proposition 5.28. □

We can now turn to the proof of Theorem 5.27.

Proof of Theorem 5.27.
Cases (a) **and** (b)**.** For (a), let us define $\Phi(L)$ for each $L \in \mathbb{N}$ by

$$\Phi(L) = \frac{\rho(L)L^2}{\psi(L)^2\log L}.$$

Let $t > 0$ and $(A_L)_{L\geq 1}$ be such that $A_L \in \Gamma_{\mathcal{A}}(L,2)^*$ for each $L \in \mathbb{N}$. Introducing the time T_L needed for the two lineages of the sample to come at distance less than $2R^B\psi(L)$, we can write

$$\begin{aligned}
&\mathbb{P}_{A_L}\left[t_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t\right]\\
&= \mathbb{P}_{A_L}\left[t_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t;\ T_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t - \Phi(L)\right] \qquad (5.40)\\
&+ \mathbb{P}_{A_L}\left[t_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t;\ T_L \leq \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t - \Phi(L)\right]. \qquad (5.41)
\end{aligned}$$

Using the strong Markov property at time T_L and the uniform convergence derived in Proposition 5.28(a), we obtain that the expression in (5.41) tends to 0 as $L \to \infty$ independently of the choice of $t > 0$ and $(A_L)_{L\in\mathbb{N}}$. For (5.40), note that

$$\begin{aligned}
&\Bigg|\, \mathbb{P}_{A_L}\left[t_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t;\ T_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t - \Phi(L)\right]\\
&\qquad - \mathbb{P}_{A_L}\left[T_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t\right]\Bigg|\\
&\leq \mathbb{P}_{A_L}\left[\frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t - \Phi(L) \leq T_L \leq \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t\right]. \qquad (5.42)
\end{aligned}$$

Since X^L as the same law as $\{\xi^L(2t), t \geq 0\}$ until the random time T_L, we can bound the quantity in (5.42) by working directly with the latter process. Let us apply Lemma 5.24 to $\big\{\psi(L)^{-1}\xi^L(2\rho(L)t), t \geq 0\big\}$, with

$$U_L = \frac{(1-\alpha)L^2\log L}{4\pi\sigma_B^2\psi(L)^2}, \quad \text{and } u_L = \frac{\Phi(L)}{2\rho(L)} = \frac{L^2}{2\psi(L)^2\log L}.$$

To this end, recall that this process evolves on the torus of size $\psi(L)^{-1}L$, and so the conditions to check are $U_L\psi(L)^2L^{-2} \to \infty$ and $u_L \leq \frac{L^2}{\psi(L)^2\sqrt{\log L}}$. Both conditions are fulfilled, and so by Lemma 5.24, the right-hand side of (5.42) is bounded by

$$\frac{C\Phi(L)\psi(L)^2}{\rho(L)L^2} = \frac{C}{\log L} \to 0 \qquad \text{as } L \to \infty.$$

Hence, coming back to (5.40), we can use the result of Proposition 5.26 and the uniformity in $t > 0$ and $(A_L)_{L\geq 1}$ of our estimates to obtain

$$\lim_{L\to\infty} \sup_{t\geq 0} \sup_{A_L \in \Gamma_{\mathcal{A}}(L,2)^*} \left| \mathbb{P}_{A_L}\left[t_L > \frac{(1-\alpha)\rho(L)L^2\log L}{2\pi\sigma_B^2\psi(L)^2}\, t\right] - e^{-t}\right| = 0.$$

The proof of (b) follows exactly the same lines, with $\Phi(L) = L^2(\log L)^{-1}$ and Lemma 5.24 applied to $\psi(L)^{-1}\xi^L(2\psi(L)^2\cdot)$.

Case (c). In contrast with the two previous cases, where coalescence in the limit is due to large events only, here the pair of lineages can coalesce only through a small event. To see this, let us define T_L^* as the first time at which the two lineages (indexed by L) come at distance less than $2R^s$ from each other, and τ_L as the first time at which at least one of them is affected by a large event while they are at distance less than $2R^B\psi(L)$ (i.e., while $X^L \in B(0, 2R^B\psi(L))$). Note that for each L, T_L^* and τ_L are stopping times with respect to the filtration $\{\mathcal{F}_t,\ t \geq 0\}$ associated to $\Pi_L^s \cup \Pi_L^B$ as we trace backwards in time. In addition, define $\tilde{T}_L^*$ as the entrance time of ξ^L into $B(0, 2R^s)$ and $\tilde{\tau}_L$ as the first time ξ^L makes a jump of size $\mathcal{O}(\psi(L))$ while it is lying in $B(0, 2R^B\psi(L))$. These two random times are stopping times with respect to the filtration

$\{\tilde{\mathcal{F}}_t,\ t \geq 0\}$ associated to ξ^L. We claim that for each $L \in \mathbb{N}$,

$$\big\{X^L(t),\ t < \tau_L \wedge T_L^*\big\} \overset{(d)}{=} \big\{\xi^L(2t),\ 2t < \tilde{\tau}_L \wedge \tilde{T}_L^*\big\}, \tag{5.43}$$

where the notation $\overset{(d)}{=}$ refers to equality in distribution. Indeed, as long as X^L has not entered $B(0, 2R^s)$ and no large event has affected it while it lay in $B(0, 2R^B\psi(L))$, coalescence events are impossible and the rates and distributions of the jumps of both processes are identical. We cannot include the terminal times in (5.43) since the values of the processes will differ if $\tau_L \wedge T_L^* = \tau_L$ and the corresponding event is a coalescence, but since X^L and ξ^L are jump processes with finite rates, we can easily see that the event $\{\tau_L \wedge T_L^* = \tau_L\}$ (resp., $\tilde{\tau}_L \wedge \tilde{T}_L^* = \tilde{\tau}_L$) is $\mathcal{F}_{(\tau_L \wedge T_L^*)-}$ (resp., $\tilde{\mathcal{F}}_{(\tilde{\tau}_L \wedge \tilde{T}_L^*)-}$)-measurable. Hence, for each $L \in \mathbb{N}$, $A = \wp_2(x_1, x_2)$ and $x := x_1 - x_2 \in \mathbb{T}(L)$, we have

$$\mathbb{P}_A\big[\tau_L < T_L^*\big] = \mathbb{P}_x\big[\tilde{\tau}_L < \tilde{T}_L^*\big]. \tag{5.44}$$

Let us now bound the right-hand side of (5.44) under the assumption that $\rho(L)^{-1}\psi(L)^4 \to \mathfrak{b} \in [0, \infty)$. The same computations as in the proof of Proposition 5.26 show that $\{\xi^L(2t), t \geq 0\}$ itself satisfies Assumption 5.21 with $\sigma_L^2 = 2\sigma_s^2 + o(1)$ as $L \to \infty$. Hence, Lemma 5.22 applied with $d_L = 2R^s$ gives us

$$\lim_{L\to\infty} \sup_{t\geq 0} \sup_{x_L \in \Gamma(L,1)} \left| \mathbb{P}_{x_L}\left[\tilde{T}_L^* > \frac{L^2 \log L}{2\pi\sigma_s^2}\, t\right] - e^{-t} \right| = 0.$$

Let $\theta_{\max} < \infty$ be such that for every $L \in \mathbb{N}$, the rate at which ξ^L makes a jump of size $\mathcal{O}(\psi(L))$ is bounded by $\theta_{\max}/\rho(L)$. Fixing $\varepsilon > 0$ and $K > 0$ such that $e^{-2\pi\sigma_s^2 K} < \varepsilon$, we have for L large enough and any sequence $(x_L)_{L\geq 1}$ such that $x_L \in \Gamma(L, 1)$ for every L:

$$\begin{aligned}
\mathbb{P}_{x_L}\big[\tilde{\tau}_L < \tilde{T}_L^*\big] &= \mathbb{P}_{x_L}\big[\tilde{\tau}_L < \tilde{T}_L^* \leq KL^2 \log L\big] + \mathbb{P}_{x_L}\big[\tilde{\tau}_L < \tilde{T}_L^*\ ;\ \tilde{T}_L^* > KL^2 \log L\big] \\
&\leq \mathbb{P}_{x_L}\big[\tilde{\tau}_L < KL^2 \log L\big] + \mathbb{P}_{x_L}\big[\tilde{T}_L^* > KL^2 \log L\big] \\
&\leq \mathbb{E}_{x_L}\left[1 - \exp\left\{-\frac{\theta_{\max}}{\rho(L)} \int_0^{KL^2 \log L} \mathbb{I}_{B(0,2R^B\psi(L))}\big(\xi^L(2s)\big) ds\right\}\right] + \varepsilon. \qquad (5.45)
\end{aligned}$$

A direct application of Lemma 5.23 shows that there exists $L_0 \in \mathbb{N}$, and $a_1, a_2 > 0$ independent of L, $(x_L)_{L\geq 1}$ and $K > 0$, such that for every $L \geq L_0$,

$$\mathbb{E}_{x_L}\left[\int_0^{KL^2 \log L} \mathbb{I}_{B(0,2R^B\psi(L))}\big(\xi^L(2s)\big) ds\right] \leq (a_1 + a_2 K)\psi(L)^2 \log L.$$

Hence, the first term on the right-hand side of (5.45) is bounded by

$$\mathbb{E}_{x_L}\left[\frac{\theta_{\max}}{\rho(L)} \int_0^{KL^2 \log L} \mathbb{I}_{B(0,2R^B\psi(L))}\big(\xi^L(2s)\big) ds\right] \leq \theta_{\max}(a_1 + a_2 K)\frac{\psi(L)^2 \log L}{\rho(L)},$$

which tends to 0 as $L \to \infty$, independently of the sequence $(x_L)_{L\geq 1}$ considered. As ε in (5.45) is arbitrary, we can conclude that

$$\lim_{L\to\infty} \sup_{x_L \in \Gamma(L,1)} \mathbb{P}_{x_L}\big[\tilde{\tau}_L < \tilde{T}_L^*\big] = 0,$$

and by (5.44), the same result holds for X^L and any sequence $(A_L)_{L\in\mathbb{N}}$ such that $A_L \in \Gamma_{\mathcal{A}}(L, 2)^*$ for every L. In words, we have obtained that with probability tending to 1, any pair of lineages starting at distance $\mathcal{O}(L)$ from each other gather at distance $2R^s$ before having a chance to coalesce through a large reproduction event. By using the same method as in (a) but this time with the result of Proposition 5.28 (b), we obtain the desired conclusion under the assumption that $\rho(L)^{-1}\psi(L)^4$ has a finite limit as $L \to \infty$.

When $\rho(L) \gg L^2 \log L$, with probability increasing to 1 no large events at all affect any of the lineages by the time they are gathered at distance $2R^s$ by small events. The result then follows from the same arguments, with ξ^L replaced by the motion of a single lineage subject to only small reproduction events. $\square$

Finally, let us comment on the cases not covered by the theorem, that is $\psi(L)^4 \gg \rho(L)$, $\rho(L)$ is of order at most $L^2 \log L$ and $\rho(L)^{-1}\psi(L)^2 \log L$ has a finite limit (possibly 0). When the latter limit is positive, from the results obtained so far coalescence events due to small and to large reproduction events occur on the same timescale and depend on the precise paths of the two lineages. Therefore, we do not expect t_L to be exponentially distributed (with a deterministic parameter). When $\rho(L)^{-1}\psi(L)^2 \log L$ tends to 0, the same reasoning as in the proof of (c) gives us that the probability that a large reproduction event causes the two lineages to coalesce before a time of order $L^2 \log L$ vanishes as $L \to \infty$. However, X^L does not satisfy the conditions of Section 5 (Assumption 5.21) as it does when the assumptions of (c) hold. Using instead $\ell^L := \psi(L)^{-1}X^L(\psi(L)^2\cdot)$, the time needed for the lineages to come at distance less than $2R^s$ translates into $T(\ell^L, 2R^s/\psi(L))$, which is not covered by Lemma 5.22 and requires estimates of the entrance time of the jump process into a ball of shrinking radius, which we have been unable to obtain.

It remains to prove Lemmas 5.29, 5.30 and 5.31.

Proof of Lemma 5.29. Let us start with the case $\rho(L) = \mathcal{O}(\psi(L)^2)$ as $L \to \infty$. The rate of decay of the probability of a long excursion is known for simple random walks and Brownian motion (see [**RR66**]), and so the proof of Proposition 5.26 suggests that we should consider the process $\hat{\ell}^L := \psi(L)^{-1}X^L(\rho(L)\cdot)$. But $\hat{\ell}^L$ here is not a Lévy process, since X^L is the difference of the locations of two lineages whose motions are not independent in $B(0, 2R^B\psi(L))$. However, it is not difficult to convince oneself that for each $y \in B(0,(7/4)R^B)^c$, the return time into $B(0,(3/2)R^B)$ of $\hat{\ell}^L$ starting at y is smaller than or equal to that of ℓ^L defined as the rescaled process $\psi(L)^{-1}\xi^L(\rho(L)\cdot)$ also starting at y. Indeed, the rate at which reproduction events affect at least one of the lineages is bounded from below by the rate at which a single lineage is affected, the distribution of the jumps of $\hat{\ell}^L$ and ℓ^L are identical outside $B(0, 2R^B)$ and inside this ball, coalescence events make it easier for $\hat{\ell}^L$ to enter $B(0,(3/2)R^B)$. Hence, we shall establish the desired bound for ℓ^L. In addition, we shall consider that ℓ^L evolves on $\mathbb{R}^2$ instead of $\mathbb{T}(L)$, since the return time here can only increase with the available space.

For each $L \in \mathbb{N}$, set $\sigma_0^L = 0$ and let $(\sigma_i^L)_{i\in\mathbb{N}}$ be the sequence of jump times of ℓ^L. Let $\rho(L)\theta_s$ (resp., θ_B) be the jump rate of ξ^L due to small events (resp., due to large events). The quantities θ_s and θ_B do not depend on L since μ^B, μ^s and the probability measures $\nu_r^{B,s}$ do not. For each $t \geq 0$, we have $\ell^L(t) = \ell^L(0) + \sum_{i:\sigma_i^L \leq t}\{\ell^L(\sigma_i^L) - \ell^L(\sigma_{i-1}^L)\}$, where $\big(\ell^L(\sigma_i^L) - \ell^L(\sigma_{i-1}^L)\big)_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables with covariance matrix $\upsilon_L\mathrm{Id}$. By the independence of Π_L^s and Π_L^B, we have

$$\begin{aligned}\upsilon_L &= \mathbb{P}_0\big[\text{small jump first}\big]\frac{c_s^2 + o(1)}{\psi(L)^2} + \mathbb{P}_0\big[\text{large jump first}\big](c_B^2 + o(1))\\ &= \frac{\theta_s}{\theta_s + \theta_B\rho(L)^{-1}}\frac{c_s^2 + o(1)}{\psi(L)^2} + \frac{\theta_B}{\theta_s\rho(L) + \theta_B}(c_B^2 + o(1)) \sim \frac{C}{\rho(L)}, \qquad \text{as } L \to \infty,\end{aligned}$$

where c_s^2 and c_B^2 are the variances (of the first coordinate) of a typical small and large jump respectively, and C is a positive constant. Let x in $B(0,4R^B) \setminus B(0,(7/4)R^B)$ and let W be a two-dimensional Brownian motion starting at x. For each $L \in \mathbb{N}$, by the Skorohod Embedding Theorem (see, e.g., [**Bil95**]) one can construct a sequence $(s_i^L)_{i\in\mathbb{N}}$ of stopping times such that the $W(s_i^L)$ have the same joint distributions as the $\ell^L(\sigma_i^L)$: for every $i \in \mathbb{N}$, conditionally on

$W(s_{i-1}^L)$, s_i^L is the first time greater than s_{i-1}^L at which W leaves $B\big(W(s_{i-1}^L), r_i^L\big)$, where r_i^L is a random variable independent of W and of $\{s_j^L, j < i\}$ having the same distribution as the length of the first jump of ℓ^L. Now, we claim that there exists $U > 0$ independent of L and x, such that each time W visits $B(0, R^B/2)$ and then leaves $B(0,(3/2)R^B)$, the probability that one of the s_i^L's falls into the corresponding period of time that W spends within $B(0,(3/2)R^B)$ is at least U. Indeed, set $T_0(W) = \breve{T}_0(W) = 0$ and define the sequences of stopping times $\{T_k(W), k \geq 1\}$ and $\{\breve{T}_k(W), k \geq 1\}$ by induction in the following manner:

$$\begin{aligned} T_k(W) &= \inf\big\{t > \breve{T}_{k-1}(W) :\ W(t) \in B(0, R^B/2)\big\}, \\ \breve{T}_k(W) &= \inf\big\{t > T_k(W) :\ W(t) \notin B(0,(3/2)R^B)\big\}. \end{aligned}$$

Then for each $k \in \mathbb{N}$, if j is the index of the last s_i^L before $T_k(W)$ and s_j^L corresponds to a small event, by construction we have $\big|W(s_j^L) - W(T_k(W))\big| < 2R^s\psi(L)^{-1}$ and so $W(s_j^L) \in B(0,(3/2)R^B)$ for L large enough. If s_j^L is due to a large event and $W(s_j^L) \notin B(0,(3/2)R^B)$, then necessarily $W(s_j^L) \in B(0,(5/2)R^B)$. But the exit point from a ball B of Brownian motion started at the centre of this ball is uniformly distributed over the boundary of B, and so one can define U as the minimum over (y,r) with $|y| \geq 3R^B/2$ and $|y| - R^B/2 < r \leq 2R^B$ of the probability that W started at y escapes $B(y,r)$, through the part of its boundary which lies within $B(0,(3/2)R^B)$. Hence, if we define for each $t \geq 0$ the random variable $N(t)$ as the maximal integer k such that $\breve{T}_k(W) \leq t$, we can write for each L

$$\mathbb{P}_x\big[q_1^L > u\big] \leq \mathbb{E}_x\Big[(1-U)^{N(s_{i(u,L)}^L)}\Big],$$

where $i(u,L) = \max\{j :\ \sigma_j^L \leq u\}$. Since N is a.s. a non-decreasing function of t, we have for any given $m \in \mathbb{R}_+$

$$\mathbb{P}_x\big[q_1^L > u\big] \leq \mathbb{E}_x\big[(1-U)^{N(mu)}\big] + \mathbb{P}_x\big[s_{i(u,L)}^L < mu\big]. \tag{5.46}$$

Now, $i(u,L)$ is the number of points of the Poisson point processes Π_L^s and Π_L^B which fall into the time interval $[0, u\rho(L)]$ on the original timescale, it is therefore a Poisson random variable with parameter $u(\theta_s\rho(L) + \theta_B)$. If $a > 0$, then by the Markov inequality

$$\mathbb{P}\big[i(u,L) \leq au\theta_s\rho(L)\big] \leq e^{au\theta_s\rho(L)}\mathbb{E}\big[e^{-i(u,L)}\big] = \exp\big\{u\theta_s\rho(L)(a+e^{-1}-1) + u\theta_B(e^{-1}-1)\big\},$$

so that this quantity converges exponentially fast to 0 for $a > 0$ small enough. On the event $\{i(u,L) > au\theta_s\rho(L)\}$, $s_{i(u,L)}^L$ is the sum of at least $au\theta_s\rho(L)$ i.i.d. random variables, with exponential moments and means $\mathbb{E}[(r_1^L)^2] = 2v_L$ (of order $\rho(L)^{-1}$), and so classical large deviation estimates gives us that for $m > 0$ small enough, $\mathbb{P}_x\big[i(u,L) > au\theta_s\rho(L),\ s_{i(u,L)}^L < mu\big]$ decreases exponentially in u. Let us now prove that

$$\mathbb{P}_x[N(mu) \leq K\log\log u] \leq C\,\frac{\log\log u}{\log u}$$

for a constant $C > 0$ independent of $x \in B(0,4R^B) \setminus B(0,(7/4)R^B)$ and u large enough (again independently of x). The reasoning is identical to that made to arrive at (5.39), with q_i^L (resp., Q_i^L) replaced by $T_i(W)$ (resp., $\breve{T}_i(W)$). Using the fact that $C_1 := \sup_{x\in B(0,R^B/2)}\mathbb{E}_x[\breve{T}_1(W)] < \infty$ and

$$\sup_{y\in B(0,4R^B)} \mathbb{P}_x\big[T_1(W) > u\big] \leq \frac{C_2}{\log u} \tag{5.47}$$

for a constant C_2 and u large enough (see Theorem 2 in [**RR66**]), we can conclude that for each $x \in B(0, 4R^B) \setminus B(0, (7/4)R^B)$, and u large enough,

$$\mathbb{P}_x\big[N(mu) \leq \log\log u\big] \leq \frac{2C_1(\log\log u)^2}{mu} + \frac{C_2 \log\log u}{\log\big(mu/(2\log\log u)\big)} \leq \frac{C' \log\log u}{\log u},$$

again for $C' > 0$ and u large enough independently of x. Coming back to (5.46), we obtain for a constant $C'' > 0$ and for all $x \in B(0, 4R^B) \setminus B(0, (7/4)R^B)$,

$$\begin{aligned} \mathbb{P}_x\big[q_1^L > u\big] &\leq (1-U)^{\log\log u} + \mathbb{P}_x\big[N(mu) \leq \log\log u\big] + e^{-C''u} \\ &\leq (1-U)^{\log\log u} + \frac{C' \log\log u}{\log u} + e^{-C''u}. \end{aligned} \tag{5.48}$$

Define $g(u)$ as the expression on the right-hand side of (5.48) to obtain the result.

When $\psi(L)^2\rho(L)^{-1} \to 0$, the probability that a large event occurs by time $u\psi(L)^2$ is given by

$$1 - \exp\bigg\{ -\theta_B u \frac{\psi(L)^2}{\rho(L)} \bigg\} \to 0 \qquad \text{as } L \to \infty.$$

On the event that no large events occur by time $u\psi(L)^2$, the first visit of W into $B(0, R^B/2)$ will produce a time s_i^L such that $W(s_i^L) \in B(0, (3/2)R^B)$ with probability 1 for the reason expounded above, and so the first term on the right-hand side of (5.46) is now the probability that $T_1(W)$ is greater than mu. The inequality in (5.47) and the exponential decay of $\mathbb{P}_x[s_{i(u,L)}^L < mu]$ now imply the result. $\square$

Proof of Lemma 5.30. The arguments are slightly different according to whether $\rho(L)\psi(L)^{-2}$ is bounded or tends to infinity as $L \to \infty$. Let us consider the first case. Recall the definition of $\rho(L)^{-1}\theta_B$ given in the proof of Lemma 5.29 as the maximal rate at which a lineage is affected by a large event. The coalescence rate of two lineages is then bounded by $2\rho(L)^{-1}\theta_B$, regardless of their locations. By our Assumption (5.35), there exist $r \in (0, R^B)$ and $\delta > 0$ such that $\mathrm{Leb}\big(\{r' \in [r, r+\delta] : \nu_r^B \notin \{\delta_0, \delta_1\}\}\big) > 0$. We shall use these events to send the two lineages at distance at least $(7/4)R^B\psi(L)$ from each other, whatever their initial separation was. The proof is quite natural, so we just give the main arguments. If only large jumps occurred, then if a sequence of at least $7R^B/(2r)$ large events increased $|X^L|$ by at least $(r/2)\psi(L)$ each before the first coalescence happened, X^L starting within $B(0, (3/2)R^B\psi(L))$ would certainly leave $B(0, (7/4)R^B\psi(L))$. Moreover, a large event affecting X^L and conditioned on not leading to a coalescence biases the jump towards increasing $|X^L|$ (we do not allow some centres that are too close to both lineages). This remark and (5.35) guarantee that the rate at which these separating events occur is bounded from below by $\rho(L)^{-1}\theta_{\mathrm{sep}}$, where θ_{sep} is a positive constant. The total rate at which large events affect X^L is bounded by $2\rho(L)^{-1}\theta_B$, and so there is a positive probability p_{sep}, independent of the starting point of X^L, that X^L leaves $B(0, (7/4)R^B\psi(L))$ before coming back to 0 through a (large) coalescence event. As regards the effect of small events, recall that we assumed that $\rho(L)\psi(L)^{-2}$ is bounded. Hence, the probability that X^L starting from $B(0, r\psi(L)/2)^c$ does not enter $B(0, 2R^s)$ after a time of order $\mathcal{O}(\rho(L))$ only through small jumps is bounded from below and by the symmetry of these small jumps, with probability at least $1/2$ the radius of X^L increases between two large jumps. Hence, up to modifying p_{sep} to take into account the effect of the small jumps, the probability that X^L leaves $B(0, (7/4)R^B\psi(L))$ before coming back to 0 is still bounded from below by $p_{\mathrm{sep}} > 0$.

By the definition of R^B and the assumption made just after this definition, large events of size close to R^B occur at a positive rate and lead to the coalescence of the lineages with positive probability, so that the waiting time for the coalescence of two lineages at distance at most $(7/4)R^B\psi(L)$

is bounded by $\rho(L)$ times an exponential with positive parameter γ. This gives us that $\rho(L)^{-1}Q_1^L$ is stochastically bounded by $\sum_{i=1}^k N_i$, where k is geometric with success probability $p_{\mathrm{sep}} > 0$ and $\{N_i, i \in \mathbb{N}\}$ is a sequence of i.i.d. $\mathrm{Exp}(\gamma)$ random variables, all of them independent of the initial value $x \in B(0,(3/2)R^B\psi(L))$ of X^L. We can therefore choose $C_Q = (\gamma p_{\mathrm{sep}})^{-1}$.

When $\rho(L)^{-1}\psi(L)^2 \to 0$, if we use the same reasoning as above there is a positive probability that a large event separates the two lineages at distance at least $r\psi(L)$, regardless of their distance just before this event. In addition, the rate of these separating events is at least equal to $\rho(L)^{-1}\theta_{\mathrm{sep}} > 0$. Between two large events, X^L does only small jumps, and as long as $X^L \notin B(0,2R^s)$, the Skorohod Embedding Theorem (see the proof of Lemma 5.29) enables us to assert that X^L will leave $B(0,(7/4)R^B\psi(L))$ in a time of order $\mathcal{O}(\psi(L)^2)$. Moreover, for $\varepsilon > 0$ small, the same argument shows that the probability that X^L leaves $B(0,(7/4)R^B\psi(L))$ before entering $B(0,\varepsilon\psi(L))$ is bounded from below by a constant $p_{\mathrm{esc}} > 0$ independent of L and of the value $y \in B(0,r\psi(L))^c$ of X^L just after the large jump described above. A fortiori, p_{esc} is also a lower bound for the probability that X^L started at y leaves $B(0,(7/4)R^B\psi(L))$ before entering $B(0,2R^s)$ only through small jumps, and so we obtain that between two large events such that the first large jump sends (or keeps) X^L out of $B(0,r\psi(L))$, X^L escapes $B(0,(7/4)R^B\psi(L))$ with probability at least p_{esc} (recall that the total rate of large events affecting at least one of the lineages is bounded by $2\theta_B\rho(L)^{-1}$ and $\rho(L) \gg \psi(L)^2$). Consequently, Q_1^L is this time stochastically bounded by $\sum_{i=1}^k N_i(L)$, where k is a geometric random variable with success probability $p_{\mathrm{esc}} > 0$ and for each $L \in \mathbb{N}$, $\{N_i(L), i \in \mathbb{N}\}$ is a sequence of i.i.d. $\mathrm{Exp}(\rho(L)^{-1}\theta_{\mathrm{sep}})$ random variables, all of them independent of the initial value $x \in B(0,(3/2)R^B\psi(L))$ of X^L. The desired result follows, with $C_Q = (\theta_{\mathrm{sep}}p_{\mathrm{esc}})^{-1}$. □

Proof of Lemma 5.31. The inequality in (5.36) can be restated as in (5.37) (the quantity inside the brackets then tends to 1), so we prove both inequalities using this form. Let θ_c be such that $\rho(L)^{-1}\theta_c$ is the minimum rate at which two lineages at distance at most $(1+\delta)R^B\psi(L)$ from each other coalesce (where $\delta > 0$ is defined at the beginning of the proof of Lemma 5.30). By the definition of R^B and Assumption (5.10), the rate at which a reproduction event of radius $r \in \left[R^B(1-\frac{\delta}{4})\psi(L), R^B\psi(L)\right]$ occurs and leads to the coalescence of the lineages does not vanish as L tends to infinity (when multiplied by $\rho(L)$), and so $\theta_c > 0$. Let us show that if $\eta > 0$ is small enough, the probability that X^L starting within $B(0,R^B\psi(L))$ does not leave $B(0,(1+\delta)R^B\psi(L))$ through only small jumps by time $\eta\psi(L)^2$ is bounded from below by a positive constant, independent of L large. The term inside the brackets in (5.37) will then come from the probability that a large event occurs before time η and the first such event leads to the coalescence of the lineages (i.e., a jump onto 0 for X^L).

Let $\eta > 0$ and $x \in B(0,R^B\psi(L))$, and let τ_B^L denote the epoch of the first large event affecting X^L. By the argument given above, the probability that X^L starting at x hits 0 before leaving $B(0,(1+\delta)R^B\psi(L))$ is bounded from below by the probability that X^L started at x stays within this ball until τ_B^L, τ_B^L is less than or equal to $\eta\psi(L)^2$ and the first large event leads to the coalescence of the lineages. Writing $\mathcal{E}_{L,\eta}$ for the event that X^L stays within $B(0,(1+\delta)R^B\psi(L))$ before τ_B^L and $\tau_B^L \leq \eta$, this probability is equal to

$$\mathbb{P}_x\big[\text{the first large event is a coalescence} \mid \mathcal{E}_{L,\eta}\big]\mathbb{P}_x\big[\mathcal{E}_{L,\eta}\big]. \tag{5.49}$$

If, for each $L \in \mathbb{N}$, $\rho(L)^{-1}E < \infty$ denotes the rate at which a single lineage on $\mathbb{T}(L)$ is affected by a large reproduction event, then the rate at which at least one of two lineages are affected is bounded by twice this quantity, and so the first probability in (5.49) is bounded from below by $\theta_c/(2E)$. Now, X^L experiences no large reproduction event before time τ_B^L, and so we can again use the equality in distribution stated in the proof of Proposition 5.28 (b) (we also keep the notation introduced there). Write t_{exit} for the first time X^L leaves $B(0,(1+\delta)R^B\psi(L))$, and $\tilde{t}_{\text{exit}}$ for the corresponding time for $\tilde{X}^L$ (which sees only small events). We have

$$\begin{aligned}\mathbb{P}_x\big[\mathcal{E}_{L,\eta}\big] &= \mathbb{P}_x\big[t_{\text{exit}} \geq \tau_B^L;\ \tau_B^L \leq \eta\psi(L)^2\big]\\ &= \mathbb{P}_x\big[\tilde{t}_{\text{exit}} \geq e(\tilde{X}^L);\ e(\tilde{X}^L) \leq \eta\psi(L)^2\big]\\ &\geq \mathbb{P}_x\big[\tilde{t}_{\text{exit}} \geq \eta\psi(L)^2;\ e(\tilde{X}^L) \leq \eta\psi(L)^2\big]\\ &= \mathbb{P}_x\big[e(\tilde{X}^L) \leq \eta\psi(L)^2 \;\big|\; \tilde{t}_{\text{exit}} \geq \eta\psi(L)^2\big]\mathbb{P}_x\big[\tilde{t}_{\text{exit}} \geq \eta\psi(L)^2\big]. \qquad (5.50)\end{aligned}$$

Since a pair of lineages is affected by a large event at rate at least $\rho(L)^{-1}E$, the first probability in (5.50) is bounded below for all $x \in B(0,R^B\psi(L))$ by

$$1 - \exp\bigg\{-\eta E\frac{\psi(L)^2}{\rho(L)}\bigg\}.$$

Now, if $\tilde{X}^L$ starts within $B(0,R^B\psi(L))$, it needs to cover a distance of at least $\delta R^B\psi(L)$ to exit $B(0,(1+\delta)R^B\psi(L))$. Furthermore, coalescence events tend to keep $\tilde{X}^L$ within $B(0,(1+\delta)R^B\psi(L))$, and so for each x, the second probability in (5.50) is larger than $\mathbb{P}_0[\hat{t}_{\text{exit}} \geq \eta\psi(L)^2]$, where $\hat{t}_{\text{exit}}$ is the exit time from $B(0,\delta R^B\psi(L))$ of the process $\{\hat{\xi}^L(2t), t \geq 0\}$ which experiences only small jumps. Decomposing this Lévy process into the sum of its jumps and applying Doob's maximal inequality to the submartingale $|\hat{\xi}^L|^2$, we obtain

$$\mathbb{P}_0\bigg[\sup_{0\leq t\leq \eta\psi(L)^2/2}|\hat{\xi}^L(2t)|^2 > \delta^2\psi(L)^2\bigg] \leq \frac{1}{\delta^2\psi(L)^2}\,\mathbb{E}_0\big[|\hat{\xi}^L(\eta\psi(L)^2)|^2\big] = \frac{2\eta\sigma_s^2}{\delta^2},$$

where the last equality comes from (5.25). Choosing $\eta > 0$ small enough so that the quantity above is less than 1, we obtain that for all $x \in B(0,R^B\psi(L))$

$$\mathbb{P}_x\big[\tilde{t}_{\text{exit}} \geq \eta\psi(L)^2\big] \geq \mathbb{P}_0[\hat{t}_{\text{exit}} \geq \eta\psi(L)^2] \geq 1 - \frac{2\eta\sigma_s^2}{\delta^2} =: \theta_4 > 0.$$

Combining the above and choosing $\theta_2 = \theta_4\theta_c/(2E)$ and $\theta_3 = \eta E$, we obtain (5.37). □

6.2. Convergence to Kingman's coalescent. To complete the proof of Theorem 5.11, we now turn to the genealogy of a finite sample, starting at distance $\mathcal{O}(L)$ from each other on $\mathbb{T}(L)$.

We can already see from our analysis for a single pair of lineages that our spatial Λ-coalescent is similar in several respects to the coalescing random walks dual to the two-dimensional voter and stepping-stone models with short-range interactions (see e.g. [**CG86, CG90**] for a study on $\mathbb{Z}^2$, and [**Cox89**] or [**ZCD05**] for examples on the torii $\mathbb{T}(L) \cap \mathbb{Z}^2$). It will therefore be no surprise that the analogy carries over to larger samples. In most of the papers cited above, the finite-dimensional distributions of the sequence of processes corresponding to the number of blocks of the ancestral partition are shown to converge when the initial distance between the lineages grows to infinity to those of a pure death process corresponding to a time-change of the number of blocks of Kingman's coalescent. In [**CG90**], more elaborate arguments yield the convergence of the finite-dimensional distributions of the unlabelled genealogical processes to those of Kingman's coalescent. Instead of adding a new instance of such proofs to the literature, we shall simply explain why the same method applies to our case. This will also enable us to prove the tightness of the unlabelled genealogical processes.

Proof of Theorem 5.11. **(i) Convergence of the finite-dimensional distributions.**

We follow here the proofs in [**CG86**] (for the number of blocks of the ancestral partition) and [**CG90**] (for the unlabelled genealogical process of a system of coalescing simple random walks on $\mathbb{Z}^2$). Notice that, since we work on the torii $\mathbb{T}(L)$, our rescaling of time differs from Cox and Griffeath's. Another significant difference is the fact that, in their model, lineages move independently of each other until the first time two of them are on the same site, upon which they coalesce instantaneously. In our setting, the movements of lineages are defined from the same Poisson point processes, and two lineages having reached a distance that enables them to coalesce can separate again without coalescing.

Despite these differences, Lemma 5.32 below shows that a key ingredient of their proof is still valid here: at the time when two lineages coalesce, the others are at distance $\mathcal{O}(L)$ from each other and from the coalescing pair. To state this result, we need some notation. Let τ_{ij} be the first time lineages i and j come within distance less than $2R^B\psi(L)$ (resp., $2R^s$) if $\rho(L) \ll \psi(L)^2 \log L$ (resp., $\rho(L) \gg \psi(L)^2 \log L$) and τ be the minimum of the τ_{ij}'s over all pairs considered. Let also τ^*_{ij} be the coalescence time of the ancestral lines of i and j, and τ^* be the minimum of the τ^*_{ij} over all lineages considered. Finally, for each i we shall denote the motion in $\mathbb{T}(L)$ of the block containing i by ξ^L_i.

LEMMA 5.32. *Under the conditions of Theorem 5.11, we have*

$$\lim_{L\to\infty} \sup_{A_L \in \Gamma_{\mathcal{A}}(L,4)^*} \mathbb{P}_{A_L}\left[\tau^* = \tau^*_{12}\ ;\ |\xi^L_1(\tau^*) - \xi^L_3(\tau^*)| \leq \frac{L}{\log L}\right] = 0, \tag{5.51}$$

$$\lim_{L\to\infty} \sup_{A_L \in \Gamma_{\mathcal{A}}(L,4)^*} \mathbb{P}_{A_L}\left[\tau^* = \tau^*_{12}\ ;\ |\xi^L_3(\tau^*) - \xi^L_4(\tau^*)| \leq \frac{L}{\log L}\right] = 0. \tag{5.52}$$

If we were considering the times τ_{ij} rather than τ^*_{ij}, Lemma 5.32 would follow from the same arguments as in [**CG86**] (see Lemma 1). Here, we have to work a bit harder and decompose the event in (5.51) into more cases. For each $L \in \mathbb{N}$, the probability in (5.51) is bounded by

$$\mathbb{P}_{A_L}\left[\tau < \frac{\varpi_L}{\sqrt{\log L}}\right] + \mathbb{P}_{A_L}\left[\tau \geq \frac{\varpi_L}{\sqrt{\log L}}\ ;\ \tau^* = \tau^*_{12}\ ;\ \tau \neq \tau_{12}\right] \tag{5.53}$$

$$+\,\mathbb{P}_{A_L}\left[\tau^* = \tau^*_{12}\ ;\ \frac{\varpi_L}{\sqrt{\log L}} \leq \tau = \tau_{12} < \tau^*_{12} - \frac{\varpi_L}{(\log L)^2}\right] \tag{5.54}$$

$$+\,\mathbb{P}_{A_L}\left[\tau^* = \tau^*_{12}; \frac{\varpi_L}{\sqrt{\log L}} \leq \tau = \tau_{12}; \tau_{12} \geq \tau^*_{12} - \frac{\varpi_L}{(\log L)^2}; \exists i \in \{1,2\}, \tau_{i3} \in (\tau_{12}, \tau^*_{12}]\right] \tag{5.55}$$

$$+\,\mathbb{P}_{A_L}\left[\tau^* = \tau^*_{12}; \frac{\varpi_L}{\sqrt{\log L}} \leq \tau = \tau_{12}; \forall i \in \{1,2\}, \tau_{i3} > \tau^*_{12}; |\xi^L_1(\tau^*) - \xi^L_3(\tau^*)| \leq \frac{L}{\log L}\right]. \tag{5.56}$$

Suppose first that $\rho(L) \ll \psi(L)^2 \log L$. The first term in (5.53) is bounded by the sum over $i \neq j \in \{1,\dots,4\}^2$ of $\mathbb{P}_{A_L}[\tau_{ij} < \varpi_L(\log L)^{-1/2}]$, which tends to 0 uniformly in A_L by Proposition 2.1 and the consistency of the genealogy. The quantity in (5.54) is bounded by

$$\mathbb{P}_{A_L}\left[\tau^*_{12} - \tau_{12} > \frac{\varpi_L}{(\log L)^2}\right],$$

which converges to zero as $L \to \infty$, uniformly in A_L (apply the strong Markov property at time τ_{12} and use (a) of Proposition 5.28). The expression in (5.55) is bounded by

$$\mathbb{P}_{A_L}\left[\frac{\varpi_L}{\sqrt{\log L}} \leq \tau = \tau_{12}\ ;\ \tau_{13} \in \left(\tau_{12}, \tau_{12} + \frac{\varpi_L}{(\log L)^2}\right]\right]$$
$$+\mathbb{P}_{A_L}\left[\frac{\varpi_L}{\sqrt{\log L}} \leq \tau = \tau_{12}\ ;\ \tau_{23} \in \left(\tau_{12}, \tau_{12} + \frac{\varpi_L}{(\log L)^2}\right]\right].$$

Applying the strong Markov property at time τ_{12} and using Lemma 5.24 with $\ell^L := \psi(L)^{-1}\{\xi^L_i - \xi^L_3\}((\psi(L)^2 \wedge \rho(L))\cdot)$ for each $i \in \{1,2\}$ (as in the proof of Theorem 5.27), we can conclude that each of the above terms tends to 0 uniformly in A_L. On the event described by (5.56), the differences $\{\xi^L_1 - \xi^L_2\}$ and $\{\xi^L_1 - \xi^L_3\}$ have the same distribution as two independent copies $\hat{\xi}^L$ and $\check{\xi}^L$ of the process ξ^L run at speed 2, and so if we write $\hat{T}_L$ (resp., $\check{T}_L$) the entrance time of

$\hat{\xi}^L$ (resp., $\check{\xi}^L$) into $B(0, 2R^B\psi(L))$, with a slight abuse of notation for the initial value to avoid multiplying the notation, (5.56) is bounded by

$$\mathbb{P}_{A_L}\bigg[\check{T}_L > \hat{T}_L \geq \frac{\varpi_L}{\sqrt{\log L}}; \big|\check{\xi}^L(\hat{T}_L)\big| \leq \frac{L}{\log L}\bigg] \leq \mathbb{P}_{A_L}\bigg[\hat{T}_L \geq \frac{\varpi_L}{\sqrt{\log L}}; \big|\check{\xi}^L(\hat{T}_L)\big| \leq \frac{L}{\log L}\bigg].$$

A straightforward application of Lemma 5.23 (b) with $\ell^L := \psi(L)^{-1}\check{\xi}^L((\rho(L) \wedge \psi(L)^2)\cdot)$ yields the uniform convergence of the last term to 0. Finally, the second term in (5.53) is bounded by the sum over all pairs $\{i,j\} \in \{1,\ldots,4\}^2$ such that $i \neq j$ and $\{i,j\} \neq \{1,2\}$ of

$$\begin{aligned}\mathbb{P}_{A_L}\bigg[\tau \geq \frac{\varpi_L}{\sqrt{\log L}}; \tau^* = \tau^*_{12}; \tau = \tau_{ij}\bigg] &\leq \mathbb{P}_{A_L}\bigg[\frac{\varpi_L}{\sqrt{\log L}} \leq \tau = \tau_{ij} < \tau^*_{12} - \frac{\varpi_L}{(\log L)^2}; \tau^*_{ij} > \tau^*_{12}\bigg] \\ &\quad + \mathbb{P}_{A_L}\bigg[\tau^* = \tau^*_{12} \geq \frac{\varpi_L}{\sqrt{\log L}}; \tau_{ij} \in \Big[\tau^*_{12} - \frac{\varpi_L}{(\log L)^2}, \tau^*_{12}\Big]\bigg].\end{aligned}$$

We can now conclude as we did for (5.54) and (5.55).

When $\rho(L) \gg \psi(L)^2 \log L$, we saw in the proof of Theorem 5.27 that with probability increasing to 1, a pair of lineages will not be affected by a large event during the periods of time when the lineages are at distance less than $2R^B\psi(L)$ from each other, until they come at distance less than $2R^s$. Consequently, we could consider the evolution of the lineages to be independent until their gathering time at distance $2R^s$. Because we are still considering a finite number of lineages, the arguments we used are applicable here again, and the proof of the last paragraph also yields (5.51) in this case. The proof of (5.52) is analogous, and is therefore omitted.

The other ingredients required to apply Cox and Griffeath's techniques are a control on the probability of 'collision' for two lineages during a short interval of time, obtained here in Lemma 5.24, and the uniform convergence of the coalescence time of two lineages, which constitutes our Theorem 5.27. With these estimates, one can obtain the limiting rates of decrease of the number of blocks of $\mathcal{A}^{L,u}$ (namely those of the number of blocks in Kingman's coalescent), and the fact that mergers are only binary as in [**CG86**]. In particular, the counterpart of their Proposition 2 here gives us that for each $n \in \mathbb{N}$,

$$\lim_{L\to\infty} \sup_{t\geq 0} \sup_{A_L \in \Gamma_{\mathcal{A}}(L,n)^*} \bigg| \mathbb{P}_{A_L}\big[|\mathcal{A}^{L,u}(t)| = n\big] - \exp\Big\{-\frac{n(n-1)}{2}\,t\Big\}\bigg| = 0, \tag{5.57}$$

which we state here because we shall need it for the case $\alpha = 1$. Note that in Proposition 2 of [**CG86**], the right-hand side of their equation gives the probability that the number of blocks is less than n, instead of equal to n as it is stated. Furthermore, in (5.57) the supremum is over $t \geq 0$ instead of $t \in [0,T]$ for some $T > 0$ (as in [**CG86**]). Our argument for this modification is the fact that the two quantities we are comparing are monotone decreasing in t.

Then, the same arguments lead to the proof that any pair of lineages is equally likely to be the first one to coalesce, as in Lemma 1 of [**CG90**]. The uniformity of the estimates obtained enables us to proceed by induction to show the uniform convergence (on a compact time-interval) of the one-dimensional distributions of $\mathcal{A}^{L,u}$ to those of $\mathcal{K}$, which translate into the uniform convergence of the finite-dimensional distributions, still on intervals of the form $[0,T]$. We refer to [**CG90**] for the complete proof of these results.

(ii) Tightness.

This follows easily from the fact that the labelled partition $\mathcal{A}^L$ with initial value in $\Gamma_{\mathcal{A}}(L,n)^*$ for some $n \in \mathbb{N}$ lies in $\Gamma_{\mathcal{A}}(L,n)$ immediately after each coalescence event, with probability tending to 1. Indeed, for each $L \in \mathbb{N}$, let $\gamma_1^L < \ldots < \gamma_{n-1}^L$ be the ranked epochs of jumps of $\mathcal{A}^{L,u}$ (if less than $n-1$ jumps occur, then the last times are equal to $+\infty$ by convention). Let also $n \in \mathbb{N}$, $A_L \in \Gamma_{\mathcal{A}}(L,n)^*$ for every $L \geq 1$, and following [**EK86**], for every $\delta, T > 0$ let

$w'(\mathcal{A}^{L,u},T,\delta)$ denote the modulus of continuity of the process $\mathcal{A}^{L,u}$ on the time interval $[0,T]$ and with time-step δ. Let $\varepsilon > 0$. With the convention that $(+\infty)-(+\infty)=+\infty$, we have

$$\mathbb{P}_{A_L}\big[w'(\mathcal{A}^L,T,\delta) > \varepsilon\big] \leq \sum_{k=2}^{n} \mathbb{P}_{A_L}\big[\gamma_k^L - \gamma_{k-1}^L < \delta\big]. \tag{5.58}$$

An easy recursion using the fact that we consider only finitely lineages and the uniform bounds obtained in Lemma 5.32 enables us to write that for all $k \in \{1,\dots,n-1\}$,

$$\sup_{A_L' \in \Gamma_{\mathcal{A}}(L,n)^*} \mathbb{P}_{A_L'}\big[\gamma_k^L < \infty\,;\, \mathcal{A}^L(\varpi_L \gamma_k^L) \notin \Gamma_{\mathcal{A}}(L,n)\big] \to 0, \qquad \text{as } L \to \infty.$$

This result and an application of the strong Markov property at time γ_{k-1}^L yield

$$\begin{aligned}\mathbb{P}_{A_L}\big[\gamma_k^L - \gamma_{k-1}^L < \delta\big] &= \mathbb{E}_{A_L}\big[\mathbb{I}_{\{\mathcal{A}^L(\varpi_L\gamma_{k-1}^L)\in\Gamma_{\mathcal{A}}(L,n)\}}\mathbb{P}_{\mathcal{A}^L(\varpi_L\gamma_{k-1}^L)}[\gamma_1^L < \delta]\big] + o(1) \\ &\leq \frac{(n-k)(n-k-1)}{2} \sup_{A_L'\in\Gamma_{\mathcal{A}}(L,2)^*} \mathbb{P}_{A_L'}[\gamma_1^L < \delta] + o(1) \end{aligned} \tag{5.59}$$

as $L \to \infty$, where the last line uses the consistency of the genealogy to bound the probability that a first coalescence event occurs to the sample of lineages before δ by the sum over all pairs of lineages of this sample of the probability that they have coalesced by time δ (note that there are at most $(n-k)(n-k-1)/2$ possible pairs just after γ_{k-1}^L). But these probabilities converge uniformly to $1-e^{-\delta}$ by Theorem 5.27, and so for δ small enough, we can make the right-hand side of (5.59) less than $\varepsilon/(n^3)$ for L large enough (n is fixed here). Coming back to (5.58), this gives us

$$\limsup_{L\to\infty} \mathbb{P}_{A_L}[w'(\mathcal{A}^L,T,\delta) > \varepsilon] \leq \varepsilon.$$

Since $\mathcal{P}_n$ is a compact metrisable space, we can apply Corollary 3.7.4 in [**EK86**] to complete the proof. □

7. Proof of Theorem 5.15

We now turn to the case $\alpha = 1$. We still have small reproduction events of size $\mathcal{O}(1)$, but now large events have sizes $\mathcal{O}(L)$ (and rate $\mathcal{O}(\rho(L)^{-1})$), so that they cover a non-negligible fraction of the torus $\mathbb{T}(L)$. By Lemma 5.23, if the lineages were only subject to small reproduction events, the location of a single lineage would be nearly uniformly distributed on $\mathbb{T}(L)$ after a time $t_L \gg L^2$. This suggests several limiting behaviours for the genealogical process $\mathcal{A}^L$, according to how $\rho(L)$ scales with L^2:

- If $\rho(L)$ is order at most $\mathcal{O}(L^2)$, then large reproduction events occur at times when the locations of the lineages are still correlated with their starting points, and so we expect space (i.e., labels in the representation we adopted) to matter in the evolution of $\mathcal{A}^L$.
- If $L^2 \ll \rho(L) \ll L^2 \log L$, then the lineages have the time to homogenise their locations over $\mathbb{T}(L)$ before the first large event occurs, but not to come at distance $2R^s$ from each other. Hence, large events should affect lineages independently of each other, and bring the genealogy down to the common ancestor of the sample before any pair of lineages experiences a coalescence due to small events.
- If $\rho(L) \approx L^2 \log L$, the fact that pairs of lineages have now the time to gather at distance $2R^s$ should add a Kingman part (i.e., almost surely binary mergers) to the genealogical process obtained in the previous point.

- If $\rho(L) \gg L^2 \log L$, Kingman's coalescent due to small reproduction events should bring the ancestry of a sample of lineages down to a single lineage before any large event occurs, so that the limiting genealogy will not see these large events.

Proof of Theorem 5.15. For (a), let us write down the generator $\overline{\mathcal{G}}_L$ of $\bar{A}^L$ applied to functions of the $\mathbb{T}(1)$-labelled partitions of $\{1, \dots, n\}$. Recall the notation x_a for the label of the block a of a labelled partition $A \in \mathcal{P}_n^\ell$ (introduced in 5.6), and write $|A|$ for the number of blocks of A. For each $L \geq 1$, f of class C^3 with respect to the labels and $A \in \mathcal{P}_n^\ell$ such that any pair (a_1, a_2) of blocks of A satisfies $|x_{a_1} - x_{a_2}| \geq (2R^s)/L$, we have

$$\begin{aligned}\overline{\mathcal{G}}_L f(A) = \rho(L) \sum_{i=1}^{|A|} \int_{\mathbb{T}(1)} dy \int_0^{R^s} \mu^s(dr) \frac{L_r(y)}{\pi r^2} \int_0^1 \nu_r^s(du) u \\ \times \Big[f\Big(A \setminus \{(a_i, x_{a_i})\} \cup \Big\{\Big(a_i, x_{a_i} + \frac{y}{L}\Big)\Big\}\Big) - f(A)\Big] + \mathcal{G}^{(B)}(A),\end{aligned} \tag{5.60}$$

where we wrote $A = \big\{(a_1, x_{a_1}), \dots, (a_{|A|}, x_{a_{|A|}})\big\}$ and

$$\begin{aligned}\mathcal{G}^{(B)}(A) = \int_{\mathbb{T}(1)} dz \int_0^{(\sqrt{2})^{-1}} \mu^B(dr) \int_{B(z,r)} \frac{dy}{V_r} \sum_{I \subset \{1,\dots,|A|\}} \prod_{i \in I} \mathbb{I}_{\{x_i \in B(z,r)\}} \prod_{j \notin I} \mathbb{I}_{\{x_j \notin B(z,r)\}} \\ \times \sum_{J \subset I} \int_0^1 u^{|J|}(1-u)^{|I|-|J|} \nu_r^B(du) \Big[f\Big(A \setminus \Big(\bigcup_{i \in J} \{(a_i, x_{a_i})\}\Big) \cup \Big\{\Big(\bigcup_{i \in J} a_i, y\Big)\Big\}\Big) - f(A)\Big]\end{aligned}$$

is the generator of the coalescence events due to large reproduction events (recall V_r is the volume of the ball $B_{\mathbb{T}(1)}(0, r)$). Note that $\mathcal{G}^{(B)}$ does not depend on L. Let us look at a particular term in the sum on the right-hand side of (5.60). Since f is of class C^3 with respect to the labels of the blocks, a Taylor expansion and the symmetry of the jumps due to small events give us

$$\begin{aligned}\rho(L) \int dy \int_0^{R^s} \mu^s(dr) \frac{L_r(y)}{\pi r^2} \int_0^1 \nu_r(du) u \Big[f\Big(A \setminus \{(a_i, x_{a_i})\} \cup \Big\{\Big(a_i, x_{a_i} + \frac{y}{L}\Big)\Big\}\Big) - f(A)\Big] \\ = \frac{\rho(L)}{L^2} \frac{\sigma_s^2}{2} \Delta_i f(A) + \mathcal{O}\Big(\frac{\rho(L)}{L^3}\Big),\end{aligned}$$

where Δ_i is the Laplacian operator on $\mathbb{T}(1)$ applied to the label of the block a_i only. Since $\rho(L) L^{-2} \to b \in [0, \infty)$ by assumption and because f is continuous on a compact space, we obtain that $\overline{\mathcal{G}}_L f$ defined on the compact set $E_L := \big\{A \in \mathcal{P}_n^\ell : L|x_{a_i} - x_{a_j}| \geq 2R^s,\ \forall i \neq j\big\}$ converges uniformly towards

$$\overline{\mathcal{G}} f(A) := \frac{b\sigma_s^2}{2} \sum_{i=1}^{|A|} \Delta_i f(A) + \mathcal{G}^{(B)} f(A).$$

Now, by the same technique as in Section 5, one can prove that the gathering time at distance $2R^s$ of two lineages starting at distance $\mathcal{O}(L)$ on $\mathbb{T}(L)$ and subject only to small events converges uniformly on the time scale $\frac{L^2 \log L}{\pi \sigma_s^2}$ to an $\mathrm{Exp}(1)$ random variable (in the sense of Lemma 5.22). In addition, since the new location of a lineage affected by a large event is chosen uniformly over a ball of $\mathbb{T}(1)$ whose radius is of order $\mathcal{O}(1)$, if a large event affects a pair of lineages but does not lead to their coalescence, then the probability that the lineages are at distance less than $(\log L)^{-1}$ just after the event vanishes as $L \to 0$. If we call $\check{T}_L^*$ the first time at which two lineages on $\mathbb{T}(L)$ are gathered at distance $2R^s$ and t_L^* their coalescence time in the original timescale, we readily obtain that for any $u > 0$, and $x_1' \neq x_2' \in \mathbb{T}(1)^2$,

$$\lim_{L \to \infty} \mathbb{P}_{\wp_2(Lx_1', Lx_2')} \big[t_L^* > \check{T}_L^* \,;\, \check{T}_L^* \leq \rho(L) u \big] = 0.$$

(Use the fact that the pair of lineages experiences a Poisson number of large reproduction events before time $\rho(L)u$). It follows that, if $u > 0$ is fixed, we can use the consistency of the genealogy and write

$$\mathbb{P}_{\wp_n(L\mathbf{x})}[\exists t \in [0,u] : \bar{\mathcal{A}}^L(t) \notin E_L] \leq \sum_{i<j=1}^{n} \mathbb{P}_{\{(\{i\},Lx_i),(\{j\},Lx_j)\}}\big[t_L^* > \check{T}_L^*; \check{T}_L^* \leq \rho(L)u\big] \to 0.$$

Consequently, one can use Corollary 4.8.7 in [**EK86**] (with E_L as the subspace of interest in Condition (f)) to conclude that the law under $\mathbb{P}_{\wp_n(L\mathbf{x})}$ of $\bar{A}^L$ converges to that of $\bar{\mathcal{A}}^{\infty,b}$ as processes in the Skorohod space of all càdlàg paths with values in the $\mathbb{T}(1)$-labelled partitions of $\{1,\ldots,n\}$.

Let us now prove (b). Recall the assumption that the total rate at which large events occur is finite, that is $M := \mu^B([0,(\sqrt{2})^{-1}]) < \infty$. Let us first analyse what happens during the first event which may affect the unlabelled ancestral partition.

Define for each $L \geq 1$ the stopping time e_1^L by the following property: $\rho(L)e_1^L$ is the first time on the original timescale at which either a large event occurs, or $\mathcal{A}^L$ undergoes a coalescence event due to small reproduction events. Since large and small reproduction events are independent, $\rho(L)e_1^L$ has the same distribution as the minimum of two following independent random times:

- the first time of occurrence of a large event, that is an $\mathrm{Exp}\big(M/\rho(L)\big)$-random variable.
- the time t_L^* at which a first coalescence event occurs between lineages of the genealogical process $\tilde{\mathcal{A}}^L$ evolving only owing to small reproduction events.

By (5.57) applied to the case $\rho(L) \equiv +\infty$ (i.e., no large events occur), $\frac{2\pi\sigma_s^2}{L^2\log L}\, t_L^*$ converges to an $\mathrm{Exp}\big(n(n-1)/2\big)$-random variable under $\mathbb{P}_{A_L}$, uniformly in $(A_L)_{L\in\mathbb{N}}$ such that $A_L \in \Gamma_{\mathcal{A}}(L,n)^*$ for every L. It is then straightforward to obtain

$$\lim_{L\to\infty} \sup_{t\geq 0} \sup_{A_L\in\Gamma_{\mathcal{A}}(L,n)^*} \left| \mathbb{P}_{A_L}[e_1^L > t] - \exp\Big(-\Big\{M + \beta\frac{n(n-1)}{2}\Big\}t\Big)\right| = 0, \tag{5.61}$$

where the formulation is also valid for $\beta = 0$. Also, by the independence of Π_L^s and Π_L^B, for every $(A_L)_{L\in\mathbb{N}}$ as above we have (with an abuse of notation)

$$\mathbb{P}_{A_L}\big[\rho(L)e_1^L = t_L^*\big] = \mathbb{E}_{A_L}\Big[\exp\Big\{-\frac{M}{\rho(L)}\,t_L^*\Big\}\Big].$$

Using Fubini's theorem and a change of variable, we can write

$$\begin{aligned}
\mathbb{E}_{A_L}\Big[\exp\Big\{-\frac{M}{\rho(L)}\,t_L^*\Big\}\Big] &= \int_0^1 \mathbb{P}_{A_L}\Big[\exp\Big\{-\frac{M}{\rho(L)}\,t_L^*\Big\} > s\Big]ds \\
&= \int_0^1 \mathbb{P}_{A_L}\Big[\frac{2\pi\sigma_s^2}{L^2\log L}\,t_L^* < -\frac{2\pi\sigma_s^2\rho(L)}{ML^2\log L}\,\log s\Big]ds \\
&= \frac{ML^2\log L}{2\pi\sigma_s^2\rho(L)}\int_0^\infty e^{-\frac{ML^2\log L}{2\pi\sigma_s^2\rho(L)}u}\,\mathbb{P}_{A_L}\Big[\frac{2\pi\sigma_s^2}{L^2\log L}\,t_L^* < u\Big]du \\
&= 1 - \frac{ML^2\log L}{2\pi\sigma_s^2\rho(L)}\int_0^\infty e^{-\frac{ML^2\log L}{2\pi\sigma_s^2\rho(L)}u}\,\mathbb{P}_{A_L}\Big[\frac{2\pi\sigma_s^2}{L^2\log L}\,t_L^* \geq u\Big]du.
\end{aligned}$$

When $\beta > 0$, we have $\frac{ML^2\log L}{2\pi\sigma_s^2\rho(L)} \to \frac{M}{\beta}$ and so we can use the uniform convergence derived in (5.57) and the fact that the distribution of t_L^* does not charge points to conclude that

$$\lim_{L\to\infty} \sup_{A_L\in\Gamma_{\mathcal{A}}(L,n)^*} \left|\mathbb{P}_{A_L}\big[\rho(L)e_1^L = t_L^*\big] - \frac{\beta n(n-1)}{\beta n(n-1) + 2M}\right| = 0.$$

The limit holds also for $\beta = 0$ by a trivial argument. A byproduct of this result is the existence of a constant $C_0 > 0$ and $L_0 \in \mathbb{N}$ such that, for all $L \geq L_0$ and $(A_L)_{L\in\mathbb{N}}$ as above, $\mathbb{P}_{A_L}[\rho(L)e_1^L < t_L^*] \geq C_0$. We shall need this fact in the next paragraph.

By Theorem 5.11 in the case $\rho(L) \equiv \infty$, up to an error term tending uniformly to 0, on the event $\{\rho(L)e_1^L = t_L^*\}$ the transition occurring to $\mathcal{A}^{L,u}$ at time $\rho(L)e_1^L$ is the coalescence of a pair of blocks, each pair having the same probability to be the one which coalesces. Let us show that, conditioned on $\{\rho(L)e_1^L < t_L^*\}$, the locations of the lineages at time $(\rho(L)e_1^L)-$ are approximately distributed as n independent uniform random variables on $\mathbb{T}(L)$. We use again the notation τ_{ij}, τ_{ij}^* and $\tau, \tau^*(= t_L^*$ here) introduced in the proof of Theorem 5.11 for the gathering time at distance $2R^s$ and the coalescence time of lineages i and j, and their minima (once again on the original timescale). These quantities depend on L, but we do not report it in the notation for the sake of clarity. In order to use our results on Lévy processes, we need to make sure that no pairs of lineages have come at distance less than $2R^s$ before time $\rho(L)e_1^L$. We have for each $L \in \mathbb{N}$

$$\mathbb{P}_{A_L}\big[\tau < \rho(L)e_1^L \big| \, \rho(L)e_1^L < t_L^*\big] \leq \sum_{i<j=1}^{n} \mathbb{P}_{A_L}\big[\tau_{ij} < \rho(L)e_1^L \big| \, \rho(L)e_1^L < t_L^*\big], \tag{5.62}$$

Each term (i,j) on the right-hand side of (5.62) is bounded by

$$\begin{aligned} &\mathbb{P}_{A_L}\big[\, \tau_{ij} < \rho(L)e_1^L - \log L \big| \, \rho(L)e_1^L < t_L^*\big] \\ &\qquad\qquad + \mathbb{P}_{A_L}\big[\rho(L)e_1^L - \log L \leq \tau_{ij} < \rho(L)e_1^L \big| \, \rho(L)e_1^L < t_L^*\big] \\ &\qquad \leq C_0^{-1}\Big\{\mathbb{P}_{A_L}\big[\tilde{\tau}_{ij}^* > \tilde{\tau}_{ij} + \log L\big] + \mathbb{P}_{A_L}\big[\tilde{\tau}_{ij} \in [\varsigma_L - \log L, \varsigma_L)\big]\Big\}, \end{aligned} \tag{5.63}$$

where for each $L \in \mathbb{N}$, ς_L is an $\mathrm{Exp}(M/\rho(L))$-random variable independent of all other variables, and $\tilde{\tau}_{ij}$ and $\tilde{\tau}_{ij}^*$ are defined as above, but for the process $\tilde{\mathcal{A}}^L$. By the strong Markov property applied at time $\tilde{\tau}_{ij}$ and the result of Proposition 5.28 (b), the first term on the right-hand side of (5.63) converges to 0 uniformly in $A_L \in \Gamma_{\mathcal{A}}(L,n)^*$. By a simple change of variable, the second term is equal to

$$M\int_0^\infty e^{-Ms}\, \mathbb{P}_{A_L}\big[\tilde{\tau}_{ij} \in [\rho(L)s - \log L, \rho(L)s)\big]ds \leq M\int_0^\infty e^{-Ms}\, C\, \frac{\log L}{L^2}\, ds \;\rightarrow\; 0,$$

where the inequality comes from Lemma 5.24. Therefore, back to (5.62) we obtain that

$$\lim_{L\to\infty} \sup_{A_L\in\Gamma_{\mathcal{A}}(L,n)^*} \mathbb{P}_{A_L}\big[\tau < \rho(L)e_1^L \big| \, \rho(L)e_1^L < t_L^*\big] = 0. \tag{5.64}$$

Now, let $D_1, \ldots, D_n$ be n measurable subsets of $\mathbb{T}(1)$, and for each $i \in \{1, \ldots, n\}$ and $L \geq 1$, let $LD_i \subset \mathbb{T}(L)$ be the dilation of D_i by a factor L. Let us show that

$$\begin{aligned} \lim_{L\to\infty} \sup_{A_L\in\Gamma_{\mathcal{A}}(L,n)^*} \bigg| \, \mathbb{P}_{A_L}\Big[\, (\xi_1^L, \ldots, \xi_n^L)(\rho(L)e_1^L-) \in (LD_1) \times \ldots \times (LD_n) \Big| \rho(L)e_1^L < t_L^*\Big] \\ - \prod_{i=1}^n \mathrm{Leb}(D_i)\bigg| = 0, \end{aligned} \tag{5.65}$$

where $\xi_i^L(t)$ denotes the location of the i-th lineage of $\mathcal{A}^L$ at time t. To do so, let us use the fact that on the event $\{\rho(L)e_1^L < t_L^*\}$, the genealogical process $\mathcal{A}^L$ up to time $\rho(L)e_1^L$ has the same distribution as $\tilde{\mathcal{A}}^L$ up to time ς_L and on the event $\{\tilde{\tau}^* > \varsigma_L\}$. We have

$$\mathbb{P}_{A_L}\Big[(\xi_1^L, \ldots, \xi_n^L)(\rho(L)e_1^L-) \in \prod_{i=1}^n (LD_i) \Big| \rho(L)e_1^L < t_L^*\Big]$$

$$\begin{aligned}
&= \frac{1}{\mathbb{P}_{A_L}[\rho(L)e_1^L < t_L^*]}\,\mathbb{P}_{A_L}\Big[(\xi_1^L,\ldots,\xi_n^L)(\rho(L)e_1^L-) \in \prod_{i=1}^n (LD_i);\ \rho(L)e_1^L < t_L^*\Big] \\
&= \frac{1}{\mathbb{P}_{A_L}[\rho(L)e_1^L < t_L^*]}\,\mathbb{P}_{A_L}\Big[(\tilde{\xi}_1^L,\ldots,\tilde{\xi}_n^L)(\varsigma_L-) \in \prod_{i=1}^n (LD_i);\ \varsigma_L < \tilde{\tau}^*\Big] \\
&= \frac{1}{\mathbb{P}_{A_L}[\rho(L)e_1^L < t_L^*]}\,\mathbb{P}_{A_L}\Big[(\tilde{\xi}_1^L,\ldots,\tilde{\xi}_n^L)(\varsigma_L-) \in \prod_{i=1}^n (LD_i);\ \varsigma_L < \tilde{\tau}\Big] + \eta_L(A_L) \\
&= \frac{M}{\mathbb{P}_{A_L}[\rho(L)e_1^L < t_L^*]} \int_0^\infty ds\, e^{-Ms}\mathbb{P}_{A_L}\Big[(\tilde{\xi}_1^L,\ldots,\tilde{\xi}_n^L)(\rho(L)s-) \in \prod_{i=1}^n (LD_i); \tilde{\tau} > \rho(L)s\Big] \\
&\qquad +\eta_L(A_L), \qquad (5.66)
\end{aligned}$$

where $\eta_L(A_L)$ tends to 0 uniformly in $(A_L)_{L\in\mathbb{N}}$ by (5.64) and the fact that $\mathbb{P}_{A_L}\big[\rho(L)e_1^L < t_L^*\big]$ does not vanish.

Let us fix $s > 0$ for a moment, and consider the corresponding probability within the integral. Up to time $\tilde{\tau}$, the movements of the lineages are distributed as n independent copies $\hat{\xi}_1^L,\ldots,\hat{\xi}_n^L$ of the motion of a single lineage, for which an easy modification of Lemma 5.23 (b) tells us that, if $(\varepsilon_L)_{L\in\mathbb{N}}$ is such that $\varepsilon_L \to 0$ but $\varepsilon_L\rho(L) \gg L^2$ as $L\to\infty$,

$$\lim_{L\to\infty}\ \sup_{v\geq\varepsilon_L}\ \sup_{x\in\mathbb{T}(L)} \Big|\mathbb{P}_x\big[\hat{\xi}^L(v\rho(L)) \in (LD)\big] - \mathrm{Leb}(D)\Big| = 0. \tag{5.67}$$

However, it is not entirely clear that this convergence will still hold for n independent lineages on the event $\{\hat{\tau} > \rho(L)s\}$ (where $\hat{\tau}$ is the first time at which at least two of them come at distance less than $2R^s$). Keeping the notation A_L for the initial value of the set of lineages and denoting the set of n (non-coalescing) motions by $\hat{\mathcal{A}}^L$, we have

$$\begin{aligned}
&\mathbb{P}_{A_L}\Big[(\hat{\xi}_1^L,\ldots,\hat{\xi}_n^L)(\rho(L)s-) \in (LD_1)\times\ldots\times(LD_n);\ \hat{\tau} \leq \rho(L)s\Big] \\
&\quad = \mathbb{E}_{A_L}\Big[\mathbb{I}_{\{\hat{\tau}\leq\rho(L)s\}}\,\mathbb{P}_{\hat{\mathcal{A}}^L(\hat{\tau})}\Big[(\hat{\xi}_1^L,\ldots,\hat{\xi}_n^L)\big((\rho(L)s-\hat{\tau})-\big) \in (LD_1)\times\ldots\times(LD_n)\Big]\Big].
\end{aligned}$$

Splitting the preceding integral into $\big\{\rho(L)(s-\varepsilon_L) \leq \hat{\tau} \leq \rho(L)s\big\}$ and $\big\{\hat{\tau} < \rho(L)(s-\varepsilon_L)\big\}$, we can use (5.67) in the latter case to write

$$\begin{aligned}
&\mathbb{E}_{A_L}\Big[\ \mathbb{I}_{\{\hat{\tau}\leq\rho(L)s\}}\,\mathbb{P}_{\hat{\mathcal{A}}^L(\hat{\tau})}\Big[(\hat{\xi}_1^L,\ldots,\hat{\xi}_n^L)\big((\rho(L)s-\hat{\tau})-\big) \in \prod_{i=1}^n(LD_i)\Big]\Big] \\
&\qquad = \mathbb{E}_{A_L}\Big[\mathbb{I}_{\{\rho(L)(s-\varepsilon_L)\leq\hat{\tau}\leq\rho(L)s\}}\,\mathbb{P}_{\hat{\mathcal{A}}^L(\hat{\tau})}\Big[(\hat{\xi}_1^L,\ldots,\hat{\xi}_n^L)\big((\rho(L)s-\hat{\tau})-\big) \in \prod_{i=1}^n(LD_i)\Big]\Big] \\
&\qquad\quad +\Big(\prod_{i=1}^n \mathrm{Leb}(D_i)\Big)\mathbb{P}_{A_L}[\hat{\tau} < \rho(L)(s-\varepsilon_L)] + \delta_L(A_L), \qquad (5.68)
\end{aligned}$$

where $(\delta_L(A_L))_{L\in\mathbb{N}}$ tends to zero uniformly in $(A_L)_{L\in\mathbb{N}}$ as L tends to infinity (we still impose that $A_L \in \Gamma_{\mathcal{A}}(L,n)^*$ for every L). By the convergence of the distribution function of $\frac{\tilde{\tau}}{L^2\log L}$ to that of an exponential random variable, uniformly in the time variable and in $(A_L)_{L\in\mathbb{N}}$, we obtain that $\mathbb{P}_{A_L}[\rho(L)(s-\varepsilon_L) \leq \hat{\tau} \leq \rho(L)s]$ converges to 0 uniformly in $(A_L)_{L\in\mathbb{N}}$ (which is also true if $\beta = 0$, i.e., $\rho(L) \ll L^2\log L$). Hence, we can find a sequence $(\delta'_L(A_L))_{L\in\mathbb{N}}$ decreasing to 0 uniformly in $(A_L)_{L\in\mathbb{N}}$, such that the whole sum on the right-hand side of (5.68) is equal to

$$\Big(\prod_{i=1}^n \mathrm{Leb}(D_i)\Big)\mathbb{P}_{A_L}[\hat{\tau} \leq \rho(L)s] + \delta'_L(A_L).$$

Likewise, we can find another sequence $(\delta''_L)_{L\in\mathbb{N}}$ decreasing to zero uniformly in $(A_L)_{L\in\mathbb{N}}$ such that

$$\mathbb{P}_{A_L}\Big[(\hat{\xi}_1^L,\dots,\hat{\xi}_n^L)(\rho(L)s-)\in(LD_1)\times\dots\times(LD_n)\Big]=\prod_{i=1}^n\mathrm{Leb}(D_i)+\delta''_L(A_L).$$

Subtracting the two last equalities, we obtain

$$\mathbb{P}_{A_L}\Big[(\hat{\xi}_1^L,\dots,\hat{\xi}_n^L)(\rho(L)s-)\in\prod_{i=1}^n(LD_i);\hat{\tau}>\rho(L)s\Big]=\Big\{\prod_{i=1}^n\mathrm{Leb}(D_i)\Big\}\mathbb{P}_{A_L}[\hat{\tau}>\rho(L)s]+o(1),$$

where the remainder decreases to 0 uniformly in $s>0$ and $(A_L)_{L\geq 1}$ such that $A_L\in\Gamma_{\mathcal{A}}(L,n)^*$ for each L. Coming back to (5.66), we obtain that it is equal to

$$\begin{aligned}\frac{M}{\mathbb{P}_{A_L}[\rho(L)e_1^L<t_L^*]}\int_0^\infty & ds\,e^{-Ms}\Big\{\Big(\prod_{i=1}^n\mathrm{Leb}(D_i)\Big)\mathbb{P}_{A_L}[\tilde{\tau}>\rho(L)s]+o(1)\Big\}\\ &=\frac{\mathbb{P}_{A_L}[\tilde{\tau}>\varsigma_L]}{\mathbb{P}_{A_L}[\tilde{\tau}^*>\varsigma_L]}\prod_{i=1}^n\mathrm{Leb}(D_i)+o(1)\\ &=\frac{\mathbb{P}_{A_L}[\tilde{\tau}^*>\varsigma_L]+o(1)}{\mathbb{P}_{A_L}[\tilde{\tau}^*>\varsigma_L]}\prod_{i=1}^n\mathrm{Leb}(D_i)+o(1),\end{aligned}$$

from which we can conclude that (5.65) holds.

Condition on the first event being a large reproduction event. By the description of such an event, the result for the genealogical process is the merger of at most one group of blocks into a bigger block. Furthermore, the transitions depend only on the number of blocks and their labels, so for convenience we derive the transition probabilities for A_L of the form $\wp_n(\mathbf{x})$ only, although we shall use the result later for more general labelled partitions. Let π be a partition of $\{1,\dots,n\}$ such that π has exactly one block of size greater than 1, which we call J. Then if the large event has centre x and radius r in $\mathbb{T}(1)$, the probability that the transition undergone by $\mathcal{A}^{L,u}$ is $\wp_n\to\pi$ is the probability that at this time, at least all the lineages in J have labels in $B(x,r)$ and are really affected by the event, and all the other lineages present in $B(x,r)$ are not affected by the event. Summing over all possible choices $I\subset\{1,\dots,n\}\setminus J$ for these 'other lineages' (I can be empty) and using (5.65), the probability of the transition $\wp_n\to\pi$ up to a vanishing error is given by

$$\begin{aligned}\sum_I V_r^{|J|+|I|}\,(1-V_r)^{n-|J|-|I|}&\int_0^1 u^{|J|}(1-u)^{|I|}\nu_r^B(du)\\ &=\int_0^1(uV_r)^{|J|}\sum_{i=0}^{n-|J|}\binom{n-|J|}{i}V_r^i(1-V_r)^{n-|J|-i}(1-u)^i\nu_r^B(du)\\ &=\int_0^1(uV_r)^{|J|}((1-u)V_r+1-V_r)^{n-|J|}\nu_r^B(du)\\ &=\int_0^1(uV_r)^{|J|}(1-uV_r)^{n-|J|}\nu_r^B(du).\end{aligned}\tag{5.69}$$

We now have the results we need to show (b). For every $L\in\mathbb{N}$, let us consider again the time e_1^L introduced earlier, and define for each integer $i\geq 2$,

$$\begin{aligned}e_i^L=\inf\big\{t>e_{i-1}^L\,:\,\rho(L)t\in\Pi_L^B\ &\text{or }\rho(L)t\text{ is the epoch of a coalescence}\\ &\text{due to small events}\big\}.\end{aligned}$$

Let us also define similar times corresponding to $\Lambda^{(\beta)}$. From the expression of its rates given in Definition 5.14, $\Lambda^{(\beta)}$ is composed of a Kingman part (i.e., only binary mergers) run at rate β, and of a set of multiple mergers due to the part $\Lambda^{(0)}$ of its Λ-measure with the atom at 0 removed. Furthermore, the finite measure $\Lambda^{(0)}$ on $[0,1]$ is given by

$$\Lambda^{(0)}(dv) = v^2 \int_0^{(\sqrt{2})^{-1}} \nu_r^B\big(\{u:\ uV_r \in dv\}\big)\mu^B(dr) = v^2 \int_0^{(\sqrt{2})^{-1}} \mathbb{I}_{\{V_r \geq v\}} \nu_r^B\Big(d\frac{v}{V_r}\Big)\mu^B(dr).$$

Following Pitman's Poissonian construction of a coalescent with multiple mergers (whose Λ-measure has no atom at 0, see [**Pit99**]), let us define Π as a Poisson point process on $\mathbb{R}_+ \times [0,1]$ with intensity $dt \otimes v^{-2}\Lambda^{(0)}(dv)$. Note that because of our assumption on M, $v^{-2}\Lambda^{(0)}(dv)$ is also a finite measure, with total mass M. The atoms of Π constitute the times at which $\Lambda^{(\beta)}$ acting on the partitions of $\mathbb{N}$ experiences a multiple collision, and the probabilities that any given lineage is affected by the event. The Kingman part of $\Lambda^{(\beta)}$ is superimposed on this construction by assigning to all pairs of blocks of the current partition independent exponential clocks with parameter β, giving the time at which the corresponding pair merges into one block.

From now on, we consider only the restriction of $\Lambda^{(\beta)}$ to $\mathcal{P}_n$, although we do not make it appear in the notation. Let e_1 be the minimum of the first time a pair of blocks of $\Lambda^{(\beta)}$ merges due to the Kingman part and of the time corresponding to the first point of Π. Define e_i in a similar manner for all $i \geq 2$, so that $(e_i)_{i\in\mathbb{N}}$ is an increasing sequence of random times at which $\Lambda^{(\beta)}$ may undergo a transition. Our goal is to show that the finite-dimensional distributions of $\big\{(e_i^L, \mathcal{A}^{L,u}(e_i^L)),\ i \in \mathbb{N}\big\}$ under $\mathbb{P}_{A_L}$ converge to those of $\big\{(e_i, \Lambda^{(\beta)}(e_i)),\ i \in \mathbb{N}\big\}$ under $\mathbb{P}_{\wp_n}$, as $L \to \infty$. Since $\mathcal{A}^{L,u}$ (resp., $\Lambda^{(\beta)}$) can jump only at the times e_i^L (resp., e_i), the fact that only finitely many jumps occur to $\Lambda^{(\beta)}$ in any compact time interval, together with Proposition 3.6.5 in [**EK86**] enable us to conclude that this convergence yields (b). We proceed by induction, by showing that for each $i \in \mathbb{N}$:

$H(i)$: if $a_L \in \Gamma_{\mathcal{A}}(L,n)$ for each L and there exists $\pi_0 \in \mathcal{P}_n$ such that for all $L \in \mathbb{N}$, $\mathrm{bl}(a_L) = \pi_0$, then as $L \to \infty$

$$\mathcal{L}_{\mathbb{P}_{a_L}}\Big(\big\{(e_1^L, \mathcal{A}^{L,u}(e_1^L)), \ldots, (e_i^L, \mathcal{A}^{L,u}(e_i^L))\big\}\Big) \Rightarrow \mathcal{L}_{\mathbb{P}_{\pi_0}}\Big(\big\{(e_1, \Lambda^{(\beta)}(e_1)), \ldots, (e_i, \Lambda^{(\beta)}(e_i))\big\}\Big).$$

(Note that a_L can have less than n blocks).

Let us start by $H(1)$. Let $t \geq 0$, $\pi \in \mathcal{P}_n$ and write n_0 for the number of blocks of π_0. We have, in the notation used in the previous paragraph (and with $\tilde{\mathcal{A}}^{L,u}$ defined as the unlabelled partition induced by $\tilde{\mathcal{A}}^L$ on the timescale $\rho(L)$),

$$\begin{aligned}
\mathbb{P}_{a_L}\big[e_1^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi\big] &= \mathbb{P}_{a_L}\big[e_1^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi;\ \rho(L)e_1^L = t_L^*\big] \\
&\quad + \mathbb{P}_{a_L}\big[e_1^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi;\ \rho(L)e_1^L < t_L^*\big] \\
&= \mathbb{P}_{a_L}\big[t_L^* \leq \rho(L)t;\ \tilde{\mathcal{A}}^{L,u}(t_L^*/\rho(L)) = \pi;\ t_L^* < \zeta_L\big] \qquad (5.70)\\
&\quad + \mathbb{P}_{a_L}\big[e_1^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi \big|\ \rho(L)e_1^L < t_L^*\big]\mathbb{P}_{a_L}\big[\rho(L)e_1^L < t_L^*\big]. \qquad (5.71)
\end{aligned}$$

By Theorem 5.11 applied with $\rho(L) \equiv +\infty$, $\tilde{\mathcal{A}}^{L,u}$ with initial value a_L converges as $L \to \infty$ to Kingman's coalescent $\mathcal{K}^{(\beta)}$ started at π_0 and run at rate β, as a process in $D_{\mathcal{P}_n}[0,\infty)$ (if $\beta = 0$, then $\tilde{\mathcal{A}}^{L,u}$ converges to the constant process equal to π_0). Hence, by the independence of $\tilde{\mathcal{A}}^L$ and ζ_L for every L and a simple time-change, the quantity in (5.70) tends to that corresponding to $\mathcal{K}^{(\beta)}$, that is

$$\mathbb{P}_{\pi_0}\big[\mathcal{K}^{(\beta)}(e_1^{\mathcal{K}}) = \pi\big]\mathbb{P}_{\pi_0}\big[e_1^{\mathcal{K}} < t \wedge \zeta\big], \qquad (5.72)$$

where $e_1^{\mathcal{K}}$ is distributed like an $\mathrm{Exp}\big(\beta\frac{n_0(n_0-1)}{2}\big)$-random variable and stands for the epoch of the first event occurring to $\mathcal{K}^{(\beta)}$, and ζ is an $\mathrm{Exp}(M)$-random variable. By the construction of $\Lambda^{(\beta)}$ given in the last paragraph, (5.72) is the probability that the first event occurring to $\Lambda^{(\beta)}$ happens before time t, is due to the Kingman part of the coalescent and leads to the transition $\pi_0 \to \pi$. For (5.71), note first that because Π_L^B and Π_L^s are independent, if we condition on $\rho(L)e_1^L$ being the time of the first point (t_1^L, x_1^L, r_1^L) of Π_L^B, then e_1^L and the pair (x_1^L, r_1^L) are independent. Hence, we have for each $L \geq 1$

$$\begin{aligned}&\mathbb{P}_{a_L}\big[e_1^L \leq t\ ;\ \mathcal{A}^{L,u}(e_1^L) = \pi \big|\ \rho(L)e_1^L < t_L^*\big]\\ &\qquad = \mathbb{P}_{a_L}\big[e_1^L \leq t \big|\ \rho(L)e_1^L < t_L^*\big]\mathbb{P}_{a_L}\big[\mathcal{A}^{L,u}(e_1^L) = \pi \big|\ \rho(L)e_1^L < t_L^*\big].\end{aligned}$$

Using (5.61) and the same reasoning as for (5.70), we can write

$$\begin{aligned}\mathbb{P}_{a_L}\big[e_1^L \leq t \big|\ \rho(L)e_1^L < t_L^*\big]\,&\mathbb{P}_{a_L}\,\big[\rho(L)e_1^L < t_L^*\big]\\ &= \mathbb{P}_{a_L}\big[e_1^L \leq t;\ \rho(L)e_1^L < t_L^*\big]\\ &= \mathbb{P}_{a_L}\big[e_1^L \leq t\big] - \mathbb{P}_{a_L}\big[e_1^L \leq t;\ \rho(L)e_1^L = t_L^*\big]\\ &\to \exp\Big\{ - \Big(M + \beta\frac{n_0(n_0-1)}{2}\Big)t\Big\} - \mathbb{P}_{\pi_0}\big[e_1^{\mathcal{K}} \leq t \wedge \zeta\big]\\ &= \mathbb{P}_{\pi_0}\big[\zeta < t \wedge e_1^{\mathcal{K}}\big],\end{aligned}$$

where the last equality comes from the fact that an $\mathrm{Exp}\big(\beta\frac{n_0(n_0-1)}{2} + M\big)$-random variable has the same distribution as the minimum of an $\mathrm{Exp}\big(\beta\frac{n_0(n_0-1)}{2}\big)$- and an $\mathrm{Exp}(M)$-random variables, independent of each other. In addition, by the calculation done in (5.69),

$$\mathbb{P}_{a_L}\big[\mathcal{A}^{L,u}(e_1^L) = \pi \big|\ e_1^L < t_L^*\big] \to \mathbb{P}_{\pi_0}\big[\Lambda^{(0)}(e_1^\Lambda) = \pi\big], \qquad \text{as } L \to \infty,$$

where e_1^Λ is the time of the first event of Π. Combining the above, and recognizing the transition probability of $\Lambda^{(\beta)}$ through the decomposition obtained, we can write

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\big[e_1^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi\big] = \mathbb{P}_{\pi_0}\big[e_1 \leq t;\ \Lambda^{(\beta)}(e_1) = \pi\big].$$

Since this result holds for each $t \geq 0$ and $\pi_0 \in \mathcal{P}_n$, using a monotone class argument we can conclude that the distribution of $\big(e_1^L, \mathcal{A}^{L,u}(e_1^L)\big)$ under $\mathbb{P}_{a_L}$ converges to the distribution of $(e_1, \Lambda^{(\beta)}(e_1))$ under $\mathbb{P}_{\pi_0}$ as $L \to \infty$. This proves $H(1)$.

Suppose that $H(i-1)$ holds for some $i \geq 2$. Let $D \subset (\mathbb{R}_+)^{i-1}$, $t \geq 0$ and $\pi_1, \ldots, \pi_i \in \mathcal{P}_n$. Let also $L \in \mathbb{N}$. By the strong Markov property applied to $\mathcal{A}^L$ at time $\rho(L)e_{i-1}^L$, we have

$$\begin{aligned}&\mathbb{P}_{a_L}\big[\big(e_1^L, \ldots, e_{i-1}^L\big) \in D;\ e_i^L - e_{i-1}^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi_1, \ldots, \mathcal{A}^{L,u}(e_i^L) = \pi_i\big]\\ &\qquad = \mathbb{E}_{a_L}\Big[\,\mathbb{I}_{\{(e_1^L,\ldots,e_{i-1}^L)\in D\}}\ \mathbb{I}_{\{\mathcal{A}^{L,u}(e_1^L)=\pi_1,\ldots,\mathcal{A}^{L,u}(e_{i-1}^L)=\pi_{i-1}\}}\\ &\qquad\qquad\qquad \times\mathbb{P}_{\mathcal{A}^L(\rho(L)e_{i-1}^L)}\big[e_1^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi_i\big]\Big].\end{aligned}$$

First, using arguments analogous to those leading to Lemma 5.32, up to an error term vanishing uniformly in $(a_L)_{L\in\mathbb{N}}$ such that $a_L \in \Gamma(L,n)$ for every $L \in \mathbb{N}$, we can consider that $\mathcal{A}^L(\rho(L)e_{i-1}^L) \in \Gamma_{\mathcal{A}}(L,n)$. As $\mathrm{bl}\big(\mathcal{A}^L(\rho(L)e_{i-1}^L)\big) = \pi_{i-1}$ for each L, we can therefore use $H(1)$ to write that

$$\lim_{L\to\infty} \mathbb{P}_{\mathcal{A}^L(\rho(L)e_{i-1}^L)}\big[e_1^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi_i\big] = \mathbb{P}_{\pi_{i-1}}\big[e_1 \leq t;\ \Lambda^{(\beta)}(e_1) = \pi_i\big],$$

and so dominated convergence and $H(i-1)$ give us

$$\lim_{L\to\infty} \mathbb{P}_{a_L}\ \big[\big(e_1^L, \ldots, e_{i-1}^L\big) \in D;\ e_i^L - e_{i-1}^L \leq t;\ \mathcal{A}^{L,u}(e_1^L) = \pi_1, \ldots, \mathcal{A}^{L,u}(e_i^L) = \pi_i\big]$$

$$
\begin{aligned}
&= \mathbb{E}_{\pi_0}\Big[\,\mathbb{I}_{\{(e_1,\dots,e_{i-1})\in D\}}\,\mathbb{I}_{\{\Lambda^{(\beta)}(e_1)=\pi_1,\dots,\Lambda^{(\beta)}(e_{i-1})=\pi_{i-1}\}}\mathbb{P}_{\pi_{i-1}}\big[e_1\leq t;\ \Lambda^{(\beta)}(e_1)=\pi_i\big]\Big]\\
&= \mathbb{P}_{\pi_0}\big[\big(e_1,\dots,e_{i-1}\big)\in D;\ e_i-e_{i-1}\leq t;\ \Lambda^{(\beta)}(e_1)=\pi_1,\dots,\Lambda^{(\beta)}(e_i)=\pi_i\big],
\end{aligned}
$$

which again yields $H(i)$ by standard arguments. The induction is now complete, and so we can conclude that the finite-dimensional distributions of the embedded Markov chain and the holding times of $\mathcal{A}^{L,u}$ under $\mathbb{P}_{a_L}$ converge as $L\rightarrow\infty$ towards those of $\Lambda^{(\beta)}$ under $\mathbb{P}_{\pi_0}$. The proof of (b) is then complete.

To finish, suppose that $\rho(L)\gg L^2\log L$. Then, we can find a sequence $\Phi(L)$ increasing to $+\infty$ such that

$$
\sup_{A\in\Gamma_{\mathcal{A}}(L,n)}\ \mathbb{P}_A[\text{ a large event affects at least one lineage before time }\Phi(L)L^2\log L]\rightarrow 0
$$

as $L\rightarrow\infty$. Hence, we can couple $\mathcal{A}^L$ with the process $\tilde{\mathcal{A}}^L$ which experiences only small events, so that the time by which they differ at step L is larger than $\Phi(L)$ with probability tending to one, uniformly in the sequence $(A_L)_{L\geq 1}$ chosen as above. By the results obtained in Section 6 with $\rho(L)\equiv+\infty$, we know that $\tilde{\mathcal{A}}^{L,u}$ converges in distribution towards $\mathcal{K}$, as a process in $D_{\mathcal{P}_n}[0,\infty)$. Since the sample size n is finite and under Kingman's coalescent, a sample of n lineages reaches a common ancestor in finite time almost surely, (c) follows. $\square$

Bibliography

[AN72] K.B. Athreya and P.E. Ney. *Branching processes*. Springer, 1972.

[BDE02] N.H. Barton, F. Depaulis and A.M. Etheridge. Neutral evolution in spatially continuous populations. *Theor. Pop. Biol.*, 61:31–48, 2002.

[BEV09] N.H. Barton, A.M. Etheridge and A. Véber. A new model for evolution in a spatial continuum. *Preprint*, 2009.

[BKE09] N.H. Barton, J. Kelleher and A.M. Etheridge. A new model for large-scale population dynamics: quantifying phylogeography. *Preprint*, 2009.

[Be09] N. Berestycki. Recent progress in coalescent theory. To appear in *Ensaios Matemáticos*, 2009.

[BEH09] N. Berestycki, A.M. Etheridge and M. Hutzenthaler. Survival, extinction and ergodicity in a spatially continuous population model. To appear in *Markov Process. Related Fields*, 2009.

[Ber96] J. Bertoin. *Lévy Processes*. Cambridge University Press, 1996.

[BL03] J. Bertoin and J.-F. Le Gall. Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields*, 126:261-288, 2003.

[Bha77] R.N. Bhattacharya. Refinements of the multidimensional central limit theorem and applications. *Ann. Probab.*, 5:1–27, 1977.

[Bil68] P. Billingsley. *Convergence of probability measures*. Wiley, 1968.

[Bil95] P. Billingsley. *Probability and Measure*. Wiley, 1995.

[B&al05] 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.*, 10:303–325, 2005.

[Can74] C. Cannings. The latent roots of certain Markov chains arising in genetics: a new approach, I.Haploid models. *Adv. Appl. Probab.*, 6:260–290, 1974.

[Ch91] B. Chauvin. Product martingales and stopping lines for branching Brownian motion. *Ann. Probab.*, 19:1195–1205, 1991.

[Cox89] J.T. Cox. Coalescing random walks and the voter model consensus times on the torus in $\mathbb{Z}^d$. *Ann. Probab.*, 17:1333–1366, 1989.

[CD02] J.T. Cox and R. Durrett. The stepping stone model: new formulas expose old myths. *Ann. Appl. Probab.*, 12:1348–1377, 2002.

[CG86] J.T. Cox and D. Griffeath. Diffusive clustering in the two-dimensional voter model. *Ann. Probab.*, 14:347–370, 1986.

[CG90] J.T. Cox and D. Griffeath. Mean field asymptotics for the planar stepping stone model. *Proc. London Math. Soc.*, 61:189–208, 1990.

[Daw93] D.A. Dawson. *Measure-valued Markov processes*. Lectures notes in Mathematics: école d'été de probabilités de Saint-Flour XXI: 1991. Springer, 1993.

[DP91] D.A. Dawson and E.A. Perkins. *Historical processes*. Memoirs Amer. Math. Soc. 454, 1991.

[DK99] P.J. Donnelly and T.G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.*, 27:166–205, 1999.

[DV75] M.D. Donsker and S.R.S. Varadhan. Asymptotics for the Wiener sausage. *Comm. Pure Appl. Math.*, 28:525–565, 1975.

[DLG02] T. Duquesne, J.-F. Le Gall. *Random trees, Lévy processes and spatial branching processes*. *Astérisque* 281. Soc. Math. France, 2002.

[Dy91a] E.B. Dynkin. A probabilistic approach to one class of nonlinear differential equations. *Probab. Th. Rel. Fields*, 89:89–115, 1991.

[Dy91b] E.B. Dynkin. Branching particle systems and superprocesses. *Ann. Probab.*, 19:1157–1194, 1991.

[Dy93] E.B. Dynkin. Superprocesses and partial differential equations. *Ann. Probab.*, 21:1185–1262, 1993.

[EW06] B. Eldon and J. Wakeley. Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics*, 172:2621–2633, 2006.

[EHR04] E. Eller, J. Hawks and J.H. Relethford. Local extinction and recolonization, species effective population size, and modern human origins. *Human Biology*, 76:689–709, 2004.

[Eng00] J. Engländer. On the volume of the supercritical super-Brownian sausage conditioned on survival. *Stochastic Process. Appl.*, 88:225–243, 2000.

[Eng07] J. Engländer. Branching diffusions, superdiffusions and random media. *Probab. Surveys*, 4:303–364, 2007.

[Eng08] J. Engländer. Quenched law of large numbers for branching Brownian motion in a random medium. *Ann. Inst. H. Poincaré Probab. Statist.*, 44:490–518, 2008.

[EdH03] J. Engländer and F. den Hollander. Survival asymptotics for branching Brownian motion in a Poissonian trap field. *Markov Process. Related Fields*, 9:363–389, 2003.

[Eth00] A.M. Etheridge. *An introduction to superprocesses*. AMS University Lecture Series, Vol. 20, 2000.

[Eth08] A.M. Etheridge. Drift, draft and structure: some mathematical models of evolution. *Banach Center Publ.*, 80:121–144, 2008.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov processes: characterization and convergence*. Wiley, 1986.

[EN80] S.N. Ethier and T. Nagylaki. Diffusion approximation of Markov chains with two time scales and applications to population genetics. *Adv. Appl. Probab.*, 12:14–49, 1980.

[Eva97] S.N. Evans. Coalescing Markov labelled partitions and a continuous sites genetics model with infinitely many types. *Ann. Inst. H. Poincaré Probab. Statist.*, 33:339–358, 1997.

[Fel51] W. Feller. Diffusion processes in genetics. *Proc. Second. Berkeley Symp.*, 227–246, 1951.

[Fels75] J. Felsenstein. A pain in the torus: some difficulties with the model of isolation by distance. *Amer. Nat.*, 109:359–368, 1975.

[Fi30] R. Fisher. *The genetical theory of natural selection*. Clarenson, Oxford, 1930.

[GdH92] A. Greven and F. den Hollander. Branching random walk in random environment: phase transitions for local and global growth rates. *Probab. Theory Related Fields*, 91:195–249, 1992.

[GLW07] A. Greven, V. Limic and A. Winter. Coalescent processes arising in a study of diffusive clustering. *Preprint*, 2007.

[Gr74] D.R. Grey. Asymptotic behaviour of continuous time, continuous state branching processes. *J. Appl. Probab.*, 11:669–677, 1974.

[IM65] K. Itô and H.P. McKean. *Diffusion processes and their sample paths.* Springer, 1965.

[JS87] J. Jacod and A.N. Shiryaev. *Limit thorems for stochastic processes*. Springer, 1987.

[Kac74] M. Kac. Probabilistic methods in some problems of scattering theory. *Rocky Mountain J. Math.*, 4:511–538, 1974.

[KL74] M. Kac and J.M. Luttinger. Bose-Einstein condensation in the presence of impurities II. *J. Math. Phys.*, 15:183–186, 1974.

[Kim53] M. Kimura. Stepping stone model of population. *Ann. Rep. Nat. Inst. Genetics Japan*, 3:62–63, 1953.

[Kin82] J.F.C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13:235–248, 1982.

[Lam67] J. Lamperti. The limit of a sequence of branching processes. *Z. Wahrsch. verw. Gebiete*, 7:271–288, 1967.

[LG] J.-F. Le Gall. Unpublished lecture notes.

[LG86] J.-F. Le Gall. Sur la saucisse de Wiener et les points multiples du mouvement brownien. *Ann. Probab.*, 14:1219–1244, 1986.

[LG88] J.-F. Le Gall. Fluctuation results for the Wiener sausage. *Ann. Probab.*, 16:991–1018, 1988.

[LG99] J.-F. Le Gall. *Spatial branching processes, random snakes and partial differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser, 1999.

[LG05] J.-F. Le Gall. Random trees and applications. *Probab. Surveys* 2:245-311, 2005.

[LS06] V. Limic and A. Sturm. The spatial Lambda-coalescent. *Electron. J. Probab.*, 11:363–393, 2006.

[Mal48] G. Malécot. *Les mathématiques de l'hérédité*. Masson et Cie, 1948.

[Möh98] M. Möhle. A convergence theorem for Markov chains arising in population genetics and the coalescent with selfing. *Adv. Appl. Probab.*, 30:493–512, 1998.

[Möh00] M. Möhle. Ancestral processes in population genetics - the coalescent. *J. Theoret. Biol.*, 204:629–638, 2000.

[Möh06] M. Möhle. On sampling distributions for coalescent processes with simultaneous multiple collisions. *Bernoulli*, 12:35–53, 2006.

[MS01] M. Möhle and S. Sagitov. A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.*, 29:1547–1562, 2001.

[Nor01] M. Nordborg. Coalescent theory, Chapter 7 in *Handbook of Statistical Genetics*. Wiley, Chichester, UK, 2001.

[NK02] M. Nordborg and S.M. Krone. Separation of time scales and convergence to the coalescent in structured populations. In *Modern Developments in Theoretical Population Genetics: The Legacy of Gustave Malécot*. Oxford University Press, Oxford, 2002.

[Not90] M. Notohara. The coalescent and the genealogical process in geographically structured populations. *J. Math. Biol.*, 29:59–75, 1990.

[PV80] G.C. Papanicolaou et S.R.S. Varadhan. Diffusion in regions with many small holes. In *Proceedings of the Vilnius Conference in Probability*. Springer, 1980.

[Per02] E.A. Perkins. *Dawson-Watanabe superprocesses and measure-valued diffusions*. Lectures notes in mathematics: école d'été de Saint-Flour XXIX:1999. Springer, 2002.

[Pit99] J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27:1870–1902, 1999.

[PS78] S.C. Port and C.J. Stone. *Brownian motion and classical potential theory*. Academic Press, New York, 1978.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Springer, 1999.

[RR66] C.J. Ridler-Rowe. On the first hitting times of some recurrent two-dimensional random walks. *Z. Wahrsch. verw. Geb.*, 5:187–201, 1966.

[Sag99] S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36:1116–1125, 1999.

[SW08] O. Sargsyan and J. Wakeley. A coalescent process with simultaneous multiple mergers for approximating the gene genealogies of many marine organisms. *Theor. Popul. Biol.*, 74:104–114, 2008.

[SF79] S. Sawyer and J. Fleischmann. The maximal geographical range of a mutant allele considered as a subtype of a Brownian branching random field. *Proc. Natl. Acad. Sci. USA*, 76:872–875, 1979.

[Sch00] J. Schweinsberg. Coalescents with simultaneous multiple collisions. *Electron. J. Probab.*, 5:1–50, 2000.

[Sch03] J. Schweinsberg. Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Process. Appl.*, 106:107–139, 2003.

[Shi88] T. Shiga. Stepping stone models in population genetics and population dynamics. *Stochastic processes in physics and engineering* (S. Albeverio et al ed.). D. Reidel Publishing Company, 1988.

[Sou84] W.P. Sousa. The role of disturbance in natural communities. *Annual Review of Ecology and Systematics*, 15:353–391, 1984.

[Spi58] F. Spitzer. Some theorems concerning two-dimensional Brownian motion. *Trans. Amer. Math. Soc.*, 87:187–197, 1958.

[Spi64] F. Spitzer. Electrostatic capacity, heat flow and Brownian motion. *Z. Wahrsch. verw. Gebiete*, 3:110–121, 1964.

[SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer, 1979.

[Szn87] A.-S. Sznitman. Some bounds and limiting results for the Wiener sausage of small radius associated to elliptic diffusions. *Stochastic Process. Appl.*, 25:1–25, 1987.

[Szn98] A.-S. Sznitman. *Brownian motion, obstacles and random media*. Springer, 1998.

[Tay09] J.E. Taylor. The genealogical consequences of fecundity variance polymorphism. *Genetics*, 182:813–837, 2009.

[TV09] J.E. Taylor and A. Véber. Coalescent processes in subdivided populations subject to recurrent mass extinctions. *Electron. J. Probab.*, 14:242–288, 2009.

[vdV98] A.W. van der Vaart. *Asymptotic statistics*. Cambridge University Press, 1998.

[Véb09] A. Véber. Quenched convergence of a sequence of superprocesses in $\mathbb{R}^d$ among Poissonian obstacles. *Stochastic Process. Appl.*, 119:2598–2624, 2009.

[Wak98] J. Wakeley. Segregating sites in Wright's island model. *Theor. Popul. Biol.*, 53:166–175, 1998.

[Wak99] J. Wakeley. Nonequilibrium migration in human history. *Genetics*, 153:1863–1871, 1999.

[Wak04] J. Wakeley. Metapopulation models for historical inference. *Mol. Ecol.*, 13:865–875, 2004.

[Wak08] J. Wakeley. *Coalescent theory. An introduction*. Roberts & Company Publishers, 2008.

[WA01] J. Wakeley and N. Aliacar. Gene genealogies in a metapopulation. *Genetics*, 159:893–905, 2001.

[Wil04] J.F. Wilkins. A separation of timescales approach to the coalescent in a continuous population. *Genetics*, 168:2227–2244, 2004.

[WW02] J.F. Wilkins and J. Wakeley. The coalescent in a continuous, finite, linear population. *Genetics*, 161:873–888, 2002.

[WH98] H.M. Wilkinson-Herbots. Genealogy and subpopulation differentiation under various models of population structure. *J. Math. Biol.*, 37:535–585, 1998.

[Wri31] S. Wright. Evolution in Mendelian populations. *Genetics*, 16:97–159, 1931.

[Wri43] S. Wright. Isolation by distance. *Genetics*, 28:114–138, 1943.

[ZCD05] I. Zähle, J.T. Cox and R. Durrett. The stepping stone model II: genealogies and the infinite sites model. *Ann. Appl. Probab.*, 15:671–699, 2005.