Electronic Journal of Probability

Vol. 1 (1996) Paper no. 6, pages 1–46.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol1/paper6.abs.html

TIME-SPACE ANALYSIS OF THE CLUSTER-FORMATION IN INTERACTING DIFFUSIONS

Klaus Fleischmann

Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstr. 39, D – 10117 Berlin, Germany
fleischmann@wias-berlin.de

Andreas Greven

Mathematisches Institut, Universität Erlangen-Nürnberg
Bismarckstr. $1\frac{1}{2}$, D – 91054 Erlangen, Germany
greven@mi.uni-erlangen.de

Abstract A countable system of linearly interacting diffusions on the interval $[0,1]$, indexed by a hierarchical group is investigated. A particular choice of the interactions guarantees that we are in the diffusive clustering regime, that is spatial clusters of components with values all close to 0 or all close to 1 grow in various different scales. We studied this phenomenon in [FG94]. In the present paper we analyze the evolution of single components and of clusters over time. First we focus on the time picture of a single component and find that components close to 0 or close to 1 at a late time have had this property for a large time of random order of magnitude, which nevertheless is small compared with the age of the system. The asymptotic distribution of the suitably scaled duration a component was close to a boundary point is calculated. Second we study the history of spatial 0- or 1-clusters by means of time scaled block averages and time-space-thinning procedures. The scaled age of a cluster is again of a random order of magnitude. Third, we construct a transformed Fisher-Wright tree, which (in the long-time limit) describes the structure of the space-time process associated with our system. All described phenomena are independent of the diffusion coefficient and occur for a large class of initial configurations (universality).

Keywords interacting diffusion, clustering, infinite particle system, delayed coalescing random walk with immigration, transformed Fisher-Wright tree, low dimensional systems, ensemble of log-coalescents

AMS subject classification Primary 60 K 35; Secondary 60 J 60, 60 J 15

Submitted to EJP on November 13, 1995. Final version accepted on April 9, 1996.

Contents

List of Figures

1 Introduction

The present paper is a second step in our program started in [FG94] to better understand the long-term behavior of interacting systems with only degenerate equilibria (i.e. steady states concentrated on traps), which typically occurs in weakly interacting (low dimensional) situations. Examples for this situation are branching models, linear voter model, linear systems in the sense of Liggett, and genetics models of the type we discuss here.

In the first step [FG94] of our study of infinite interacting systems of diffusions in $[0,1]$ in the regime of diffusive clustering we already obtained a detailed picture about the growth of clusters in space observed at single time points which get large. Furthermore we got a first rough insight in the time behavior of the process observed at a fixed finite collection of components. The aim of this second step of the program is to develop a suitable scheme which enables us to deepen the understanding of the large scale correlation structure in time *and* space.

The *purpose of the present paper* is threefold:

(i) A refinement in the analysis of the time structure of the component process which will reveal that the times spent close to the boundaries of $[0,1]$ are diverging at a *random order of magnitude* as the observation time point gets large. That is, they are of the form $\big(T(t)\big)^{\alpha}$ for suitable time scales $T(t)$ and a random variable α whose distribution will be identified. This order α is less than one, i.e. the age of the cluster is small compared to the system age.

(ii) To understand the *history* of spatial clusters. In particular, we want to relate the time a component was close to 0 or 1 to the spatial extension of the related cluster.

(iii) To construct an object which contains the information about the time-space structure of the system on a "macroscopic" scale. We call this object a *transformed Fisher-Wright tree.*

As in [FG94], another important aspect is that the results are proved for a whole class of models (*universality*) allowing quite general diffusion coefficients and initial laws. In particular, the role of the transformed Fisher-Wright tree is not restricted to only interacting Fisher-Wright diffusions. The interaction term considered here corresponds to the $d = 2$ case in usual lattice models (whereas the equivalent to the $d = 1$ case behaves again different, compare Klenke [Kle95]; see also Evans and Fleischmann [EF96]).

The phenomenon of clustering has been addressed for low dimensional branching systems; see for instance Iscoe [Isc86] (assertions on the finiteness of the total occupation times), Cox and Griffeath [CG85] (random ergodic limit in the critical dimension), and Dawson and Fleischmann [DF88] (scaling limit of time-space clumps in subcritical dimensions). For the voter model which exhibits qualitatively similar behavior as the interacting diffusion, the phenomenon of clustering has been approached in Cox and Griffeath [CG83] by studying occupation times. In the present paper we will follow a different, more direct approach for interacting diffusions. It is actually not too hard to use our results to study similar questions for the voter model on a hierarchical group.

The analysis of clusters and their evolution in time, presented in §§ 1.3–1.5 below, will proceed by viewing single components in suitable time scales (Theorem 1), large spatial averages in various time scales (Theorem 2), and time-space thinned-out systems (Theorem 3). The transformed Fisher-Wright tree is defined in § 1.2.

1.1 Model of interacting diffusions

We start by introducing the model under consideration. For a discussion of the population genetics motivation for this model we refer to [FG94] and references therein (see also Remark 1.5 below).

Definition 1.1 (interacting diffusion) Let $X = \{X_\xi(t);\ \xi \in \Xi,\ t \geq 0\}$ denote the *interacting diffusion* on the hierarchical group Ξ with fixed drift parameters $\{c_k;\ k \geq 1\}$ and diffusion coefficient g. This process is defined as follows:

For each specified initial state in $[0,1]^\Xi$, consider the *unique strong solution* X (Shiga and Shimizu [SS80]) of the following system of stochastic differential equations:

$$\mathrm{d}X_\xi(t) = \left(\sum_{k=1}^{\infty} \frac{c_k}{N^k}\left[X_{\xi,k}(t) - X_\xi(t)\right]\right)\mathrm{d}t + \sqrt{g\big(X_\xi(t)\big)}\,\mathrm{d}w_\xi(t), \quad \xi \in \Xi. \tag{1}$$

The ingredients occurring in this equation are the following:

(a) **(hierarchical group)** Ξ denotes the collection of sequences $\xi = [\xi_1, \xi_2, ...]$ with coordinates ξ_i in the finite set $\{0, ..., N-1\}$ (with $N \geq 2$ fixed), which are different from 0 only finitely often. Moreover,

$$\|\xi\| := \max\{i;\ \xi_i \neq 0\}, \qquad \text{to be read as } 0 \text{ if } \xi = 0 = [0,0,...], \tag{2}$$

denotes the "discrete norm" of ξ. Finally, Ξ is an *Abelian group* by defining the addition coordinate wise modulo N, and $\|\xi - \zeta\|$ is the "*hierarchical distance*" of ξ and ζ.

(b) **(ball average)** $X_{\xi,k}$ refers to the *empirical mean* (*ball average*) in a $k-$"ball":

$$X_{\xi,k}(t) := \tfrac{1}{N^k}\textstyle\sum_\zeta \mathbf{1}\{\|\xi - \zeta\| \leq k\}\, X_\zeta(t). \tag{3}$$

(c) **(driving Brownian motions)** $w = \{w_\xi(t);\ \xi \in \Xi,\ t \geq 0\}$ is a system of *independent standard Brownian motions* in R.

(d) **(diffusion coefficient)** g belongs to the set $\mathcal{G}^0$ of functions $g : [0,1] \longmapsto R_+$ (see Figure 1) which are *Lipschitz continuous* and satisfy

$$g(0) = 0 = g(1) \qquad \text{and} \qquad g > 0 \quad \text{on} \quad (0,1). \tag{4}$$

(e) **(drift parameters)** The sequence $\{c_k \geq 0;\ k \geq 1\}$ of drift parameters with values in R_+ satisfies $\sum_k c_k/N^{2k} < \infty$. ◇

Definition 1.2 (initial state) We often use a *random initial state* $X(0)$. Then $X(0)$ is assumed to be *independent* of the system w of driving Brownian motions. The law $\mathcal{L}\big(X(0)\big)$ is denoted by μ. In most cases we assume that μ belongs to the set $\mathcal{T}_\theta$ of all those distributions on $[0,1]^\Xi$ which are *shift ergodic* with *density* $\theta \in (0,1)$, that is $\int \mu(\mathrm{d}x)\ x_\xi \equiv \theta$. Write $\mathbb{P}_\mu := \mathbb{P}_\mu^g$ for the distribution of X if μ is the *initial law* $\mathcal{L}\big(X(0)\big)$, and $\mathbb{P}_z := \mathbb{P}_z^g$ in the degenerate case $\mu = \delta_z$, $z \in [0,1]^\Xi$. ◇

Example 1.3 (interacting Fisher-Wright) The standard example is the *interacting Fisher-Wright diffusion with diffusion parameter* b where by definition

$$g(r) := bf(r) \quad \text{with} \quad f(r) := r(1-r) \quad \text{and} \quad b > 0; \tag{5}$$

see Figure 1. By an abuse of notation, in this case we also write $\mathbb{P}_\mu^b$ and $\mathbb{P}_z^b$ for the laws of X. ◇

Example 1.4 (Ohta-Kimura) Another important case in genetics is the *interacting Ohta-Kimura diffusion* where $g = bf^2$. ◇

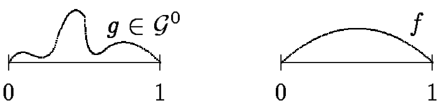

Figure 1: Diffusion coefficients: General g and standard Fisher-Wright f

Remark 1.5 (hierarchical group Ξ) We recall the following *interpretation* of Ξ used in mathematical biology: $\xi = [\xi_1, \xi_2, ...]$ *labels* the ξ_1-st member in a family, which is the ξ_2-nd family of a clan, ..., which is the ξ_k-th member of a k*–level set.* $\|\xi - \zeta\| = k$ refers to *relatives* ξ, ζ *of degree* k.

The system (1) occurs as diffusion limit of genetics models with resampling and migration (Moran model). Then X_ξ can be interpreted as a *gene frequency* of the ξ-th component (colony) of the system. (See Sawyer and Felsenstein [SF83] or Chapter 10 in Ethier and Kurtz [EK86].) ◇

The basic features of the model stem from a *competition* between drift and noise. Namely, set for the moment $c_k \equiv 0$, then all components X_ξ fluctuate *independently* according to *diffusions with coefficient* g. For instance in the Fisher-Wright case (5), X_ξ will be trapped at 0 or 1 in finite time, as indicated in Figure 2.

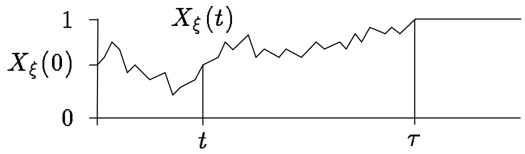

Figure 2: Under $c_k \equiv 0$: A single Fisher-Wright diffusion path trapped at 1

On the other hand, if we set $g = 0$, then X solves an infinite system of ordinary differential equations, which has the property that, under $c_k > 0$, for initial states in $\mathcal{T}_\theta$ (Definition 1.2) the solution X_t converges as $t \to \infty$ to the *constant state* $\underline{\theta}$ that is $\underline{\theta}_\xi \equiv \theta$.

Therefore in the case $c_k \not\equiv 0$, $g \neq 0$, the drift term is in competition with the basic feature of the diffusion of the single components to get trapped at $\{0, 1\}$ and in fact prevents it from getting trapped at all in finite time (except if X starts already with either of the traps $\underline{0}$ or $\underline{1}$).

In the sequel we shall study only the case where $c_k \equiv a > 0$, which is the prototype displaying a specific "critical" behavior. Namely, this special choice of the drift parameter implies first of all that we are in the regime of *clustering* (for which $\sum_k c_k^{-1} = \infty$ would suffice), that is

$$\mathbb{P}\{|X_\xi(t) - X_\zeta(t)| \geq \varepsilon\} \underset{t\to\infty}{\longrightarrow} 0, \qquad \xi, \zeta \in \Xi, \quad \varepsilon > 0. \tag{6}$$

Even more, the whole system X is in the regime of the so-called *diffusive* clustering, that is, the logarithm of the volume of clusters of neighboring components with values all close to 0 or close to 1 grow at a random linear speed if we observe the process in a suitable, in our case at an exponential, time scale; see Fleischmann and Greven [FG94]. Such behavior occurs typically if $\sum_{k=1}^n c_k^{-1}$ diverges but not exponentially fast as $n \to \infty$, while the case of $c_k^{-1} = c^{-k}$ with $c < 1$ gives different behavior, see Klenke [Kle95]. In order to keep notation reasonable we focus on $c_k \equiv a > 0$ rather than putting conditions on $\sum_{k=1}^n c_k^{-1}$. The above dichotomy is analogous to the $d = 2$, $d = 1$ dichotomy in usual lattice models as voter model, branching random walk, generalized potlatch and smoothing etc. For the voter model on $\mathbb{Z}^2$, see Cox and Griffeath [CG86]. Concerning the ergodic theory for general drift parameters c_k we refer to Cox and Greven [CG94]. For a description of the cluster-formation for general c_k, see Dawson and Greven [DG93] (concerning the mean field limit) and Klenke [Kle95]. Much of the scheme we derive to study the cluster-formation in time can be performed as well for the label set $\mathbb{Z}^2$.

1.2 Transformed Fisher-Wright tree

In order to discuss clustering phenomena, we want to introduce some objects which shall play a basic role in the description of the *genealogy of clusters* in this type of interacting models, as the backward tree in spatial branching theory does. Start with the following basic ingredients.

Definition 1.6 (Fisher-Wright) Fix $\theta \in (0,1)$.

(a) (standard Fisher-Wright diffusion Y^θ) Let B denote a standard Brownian motion, and $Y^\theta = \{Y^\theta(t);\ 0 \le t \le \infty\}$ the strong solution of

$$\mathrm{d}Y^\theta(t) = \sqrt{Y^\theta(t)\big[1 - Y^\theta(t)\big]}\,\mathrm{d}B(t), \quad 0 < t < \infty, \qquad Y^\theta(0) = \theta. \tag{7}$$

(b) (fluctuation times) We call the hitting time

$$\tau := \inf\Big\{t > 0;\ Y^\theta(t) \in \{0,1\}\Big\} \in (0,\infty) \tag{8}$$

of the traps the *fluctuation time* of Y^θ (cf. Figure 2 above).

(c) (transformed Fisher-Wright diffusion $\widetilde{Y^\theta}$) Set

$$\widetilde{Y^\theta}(\beta) := Y^\theta\big(\log(1/\beta)\big), \qquad 0 \le \beta \le 1, \tag{9}$$

and denote the *marginal laws* of this time-inhomogeneous Markov process by

$$\widetilde{Q^\theta}_\beta := \mathcal{L}\big(\widetilde{Y^\theta}(\beta)\big), \qquad 0 \le \beta \le 1. \tag{10}$$

(d) (holding time of $\widetilde{Y^\theta}$) Introduce the *holding time* of $\widetilde{Y^\theta}$:

$$\widetilde{\tau} := \mathrm{e}^{-\tau} \in (0,1). \qquad \diamond$$

Consequently, a path of the transformed Fisher-Wright diffusion $\widetilde{Y^\theta}$ starts at 0 or 1, namely with the law of $Y^\theta(\tau)$, that is with

$$(1-\theta)\delta_0 + \theta\delta_1, \tag{11}$$

stays there for the random time $\widetilde{\tau} \in (0,1)$. After $\widetilde{\tau}$, the path fluctuates like a standard Fisher-Wright diffusion but with time reversed and on a logarithmic scale, and finally ends up at time $\beta = 1$ at the deterministic value θ. (Read Figure 2 backwards.) Note that (9) can alternatively be written as $\widetilde{Y^\theta}(\mathrm{e}^{-t}) = Y^\theta(t),\ 0 \le t \le \infty$.

Next we compose a whole tree of Fisher-Wright diffusions (see Figure 3):

Definition 1.7 (Fisher-Wright tree $\mathbf{Y}^\theta$) Fix $\theta \in (0,1),\ k \ge 1$, and (deterministic) time points $0 \le s_k < \cdots < s_1 < s_0 := \infty$.

(a) (trunk) First we introduce the *trunk* of the tree denoted by $\mathbf{Y}^\theta_\infty$. It is nothing else than Y^θ from Definition 1.6 (a).

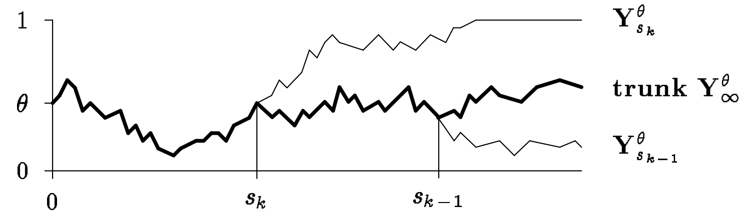

Figure 3: Fisher-Wright tree (only one branch trapped so far)

(b) **(branches)** Next we define the *branches* of the tree. Given the trunk $\mathbf{Y}^\theta_\infty$, let a branch $\mathbf{Y}^\theta_{s_k}$ split away from the trunk at the time s_k. The branch is again assumed to be standard Fisher-Wright, but defined on the time interval $[s_k, \infty]$, that is starting at time s_k with $\mathbf{Y}^\theta_{s_k}(s_k) = \mathbf{Y}^\theta_\infty(s_k)$. Proceed with the other s_i accordingly. The branches $\mathbf{Y}^\theta_{s_i}$ leave *only* from the trunk $\mathbf{Y}^\theta_\infty$ and are constructed independently of each other, given the trunk. Hence, by definition all the branches $\mathbf{Y}^\theta_{s_i}$, $k \geq i \geq 1$, are *conditionally independent* given $\mathbf{Y}^\theta_\infty$. Note that all the finitely many branches and the trunk end up in the set $\{0,1\}$ of traps after finite times.

(c) **(fluctuation time of the trunk)** As in Definition 1.6 (b), denote by τ the *fluctuation time* of the trunk. Of course, given $\mathbf{Y}^\theta_\infty(\tau) = \partial \in \{0,1\}$, all branches $\mathbf{Y}^\theta_{s_i}$ with $s_i \geq \tau$ are trapped at ∂.

(d) **(law and filtration)** For the fixed $s_k, ..., s_1$, write $\mathbf{P}^\theta$ for the law of the Fisher-Wright tree $\mathbf{Y}^\theta$ and $\{\mathcal{F}(t);\ t \geq 0\}$ for the related filtration (with $\mathcal{F}(t)$ describing the behavior of $\mathbf{Y}^\theta$ in $[0,t]$). ◇

Remark 1.8 The somewhat unexpected index $\infty = s_0$ (instead of 0 or s_{k+1}) on the symbol $\mathbf{Y}^\theta_\infty$ for the trunk of the tree will become clear below when we switch to a transformed tree. This also indicates that one could read the trunk in backward direction while then the branches, starting with $\mathbf{Y}^\theta_{s_1}$, split off in time viewed forward. This is (for good reason) the same as with the backward tree in branching theory, see for instance Chapter 12 in Dawson [Daw93]. – Note also that for typographical simplification we do not display the time points $s_k, ..., s_1$ in the notation of $\mathbf{Y}^\theta$ or $\mathbf{P}^\theta$. ◇

In analogy with the transformed Fisher-Wright diffusion $\widetilde{Y^\theta}$ defined in (9), we will introduce a transformed Fisher-Wright tree $\widetilde{\mathbf{Y}^\theta}$, see Figure 4, by starting from the Fisher-Wright tree $\mathbf{Y}^\theta$ of Definition 1.7 and switching to the time scale $\mathrm{e}^{-s} =: \beta \in [0,1]$. (Read the flat parts in Figure 4 as being exactly at the boundaries.)

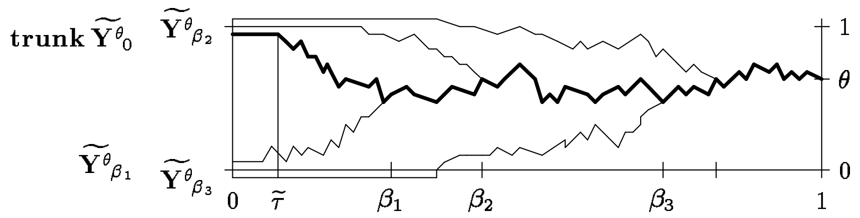

Figure 4: Transformed Fisher-Wright tree

Definition 1.9 (transformed Fisher-Wright tree $\widetilde{\mathbf{Y}^\theta}$) Fix $\theta \in (0,1)$ and $k \geq 1$, as well as $0 =: \beta_0 < \cdots < \beta_k \leq 1$.

(a) **(trunk)** The *trunk* of the transformed Fisher-Wright tree is

$$\widetilde{\mathbf{Y}^\theta}_0 := \mathbf{Y}^\theta_\infty\big(\log(1/\cdot)\big) = Y^\theta\big(\log(1/\cdot)\big) = \widetilde{Y^\theta}.$$

(b) **(branches)** The *branches* $\widetilde{\mathbf{Y}^\theta}_{\beta_1}, \ldots, \widetilde{\mathbf{Y}^\theta}_{\beta_k}$ are defined by

$$\widetilde{\mathbf{Y}^\theta}_{\beta_i}(\alpha) := \mathbf{Y}^\theta_{\log(1/\beta_i)}\big(\log(1/\alpha)\big), \qquad 0 \le \alpha \le \beta_i\,, \quad 1 \le i \le k. \tag{12}$$

Since $\alpha = 0$ is included, the trunk and all branches start from the traps $\{0,1\}$ and stay there for a positive time. The branch $\widetilde{\mathbf{Y}^\theta}_{\beta_i}$ terminates at the (deterministic) time β_i when it coalesces with the trunk; hence $\widetilde{\mathbf{Y}^\theta}_{\beta_i}(\beta_i) = \widetilde{\mathbf{Y}^\theta}_0(\beta_i)$. Consequently, $\widetilde{\mathbf{Y}^\theta}$ can be considered as a *coalescing ensemble of transformed Fisher-Wright diffusions.* Note that by Definition 1.7 (b), all the branches $\widetilde{\mathbf{Y}^\theta}_{\beta_i}$, $1 \le i \le k$, are *conditionally independent* given $\widetilde{\mathbf{Y}^\theta}_0$.

(c) **(holding time of the trunk)** As in Definition 1.6 (d), denote by $\widetilde{\tau}$ the *holding time* of the trunk $\widetilde{\mathbf{Y}^\theta}_0$. Note that

$$\widetilde{\mathbf{Y}^\theta}_{\beta_i}(\alpha) = \widetilde{\mathbf{Y}^\theta}_0(\alpha) \quad \text{if} \quad 0 \le \alpha \le \beta_i \le \widetilde{\tau}, \tag{13}$$

i.e. branches with terminal time bounded by $\widetilde{\tau}$ are trapped at the trunk.

(d) **(filtration)** Set $\widetilde{\mathcal{F}}(\beta) := \mathcal{F}\big(\log(1/\beta)\big)$, $0 \le \beta \le 1$, (with $\widetilde{\mathcal{F}}(\beta)$ describing the behavior of $\widetilde{\mathbf{Y}^\theta}$ in $[\beta, 1]$). ◇

Remark 1.10 (Trees of transformed diffusions) Trees of transformed diffusions turned out to be a basic object entering into the description of cluster-formation of interacting diffusions. This applies for instance to interacting *critical Ornstein-Uhlenbeck diffusions,* that is $g(r) = \sigma^2 > 0$, $r \in R$, where one has a tree of Brownian motions, and to *super-random walks,* that is $g(r) = \sigma^2 r$, $r \ge 0$, which lead to a tree of more complicated but explicitly known time-inhomogeneous diffusions. ◇

1.3 Time structure of components

`Expected phenomena` We start by discussing the phenomena to be described. The formal set-up and related results will be contained in the next two subsections.

For the remainder of the introduction we require:

Assumption 1.11 (initial state) X starts off with a shift ergodic distribution μ with fixed density $\theta \in (0,1)$, that is $\mu \in \mathcal{T}_\theta$. ◇

The *basic theorem* for the interacting diffusion in the regime of clustering is

$$\mathcal{L}\big(X(t)\big) \underset{t\to\infty}{\Longrightarrow} (1-\theta)\delta_{\underline{0}} + \theta\delta_{\underline{1}}\,, \qquad \text{for} \;\; \mathcal{L}\big(X(0)\big) = \mu \in \mathcal{T}_\theta\,, \tag{14}$$

(where the symbol $\Longrightarrow$ refers to weak convergence); see Cox and Greven [CG94].

Nevertheless, if we fix a label ξ in the hierarchical group Ξ, we proved in [FG94, Theorem 5] that for the corresponding *component process* $\{X_\xi(t);\, t \ge 0\}$ in $[0,1]$,

$$\limsup_{t\to\infty} X_\xi(t) = 1 \quad \text{and} \quad \liminf_{t\to\infty} X_\xi(t) = 0 \quad \text{a.s.} \tag{15}$$

In this sense, opposed to a system of independent diffusions, each component X_ξ *oscillates* "between both traps" infinitely often. As a rule, it actually even spends asymptotically fraction one of the time close to the traps $\{0,1\}$, [FG94, Theorem 4]. (Note that the oscillation property (15) can be interpreted as a type of *"recurrence"* property for clusters.)

We now want to know more about the durations for which X_ξ is close to 1 or close to 0 (*life times of clusters*, or alternatively about correlation length in the time of our system). This should be closely related to the spatial cluster-formation.

We studied the cluster extensions in space in [FG94, Theorem 3]: At time $N^{\beta t}$ (as $t \to \infty$), the spatial clusters are of "size" αt, where α is a *random* element of the open interval $(0, \beta)$ (with $\beta > 0$ fixed which could be set to 1 by scaling). More precisely, we have $\alpha = \beta\,\widetilde{\tau}$ with $\widetilde{\tau}$ the holding time of the transformed Fisher-Wright diffusion (Definition 1.6 (d)).

Or from another point of view, at time scale $N^{\beta t}$ correlations in space are built within distances of order αt, with the same random α. Or turned around, clusters of a spatial extension over a ball of radius αt need at least time $N^{(\alpha+\varepsilon)t}$ to be formed with positive probability, for some $\varepsilon > 0$. Combined with (15) this means that smaller clusters keep being overturned or melted with other smaller ones. This suggests that in order to describe the sequence of *holding times* of values close to 1 or close to 0 on a large scale, we should encounter four interesting phenomena. Namely the holding times should be

- of a *random* order of magnitude,
- asymptotically *small* compared with the age of the system and
- (stochastically) *monotone* and of increasing order of magnitude, within the correlation length,
- comparable with the correlation length.

To see that the correlation length is of a smaller order than the system age, look (for the fixed ξ) at $\{X_\xi(\beta t);\ 0 < \beta \le 1\}$. The law of this process converges as $t \to \infty$ to a *"stationary"* 0-1*–valued noise*; see Proposition 6.1 at p. 39. (Compare this phenomenon of a noisy behavior in time with the occurrence of a spatial "isolated Poissonian noise" in the analysis of the clumping in the time-space picture for branching systems in low dimensions [DF88].)

To elaborate on this point look at the exponential scale $N^{\beta t}$ and at a *component process* $\{X_\xi(N^{\beta t});\ 0 < \beta \le 1\}$ (for the fixed ξ). As $t \to \infty$ we get the same limiting "stationary" 0-1–noise. That is, the limit is independent in each "macroscopic" time point β, where the common one-dimensional marginal is just (11), with $\theta \in (0,1)$ the initial density of the system. The latter fact follows from (14).

This indicates that in order to capture time correlations we have to study X_ξ after very long times but on a much finer scale than β as it appears in $N^{\beta t}$. To accomplish that, we will look *backwards* from late time points N^T in time scales of smaller order. This will be incorporated formally by the following set-up.

`Results` To capture the structure of the correlations in time of the component process, we will look at an asymptotically small neighborhood of a late time point. For fixed $\xi \in \Xi$ and $T > 0$, we define the *scaled component process,*

$$U^T_\beta := X_\xi\big(N^T - N^{\beta T}\big), \qquad 0 \le \beta < 1, \tag{16}$$

that is $\beta \in [0,1)$ becomes the "macroscopic backward time". Consequently, from the "terminal time" N^T we look *backwards* for the amount $N^{\beta T}$ where β varies in $[0,1)$. Note that $N^T - N^{\beta T} \sim N^T$ as $T \to \infty$, so that the whole process U^T indeed describes the behavior "close to" N^T.

Recall that $g \in \mathcal{G}^0$ and $\mu \in \mathcal{T}_\theta$ with $\theta \in (0,1)$. We denote by $\stackrel{\text{fdd}}{\Longrightarrow}$ weak convergence of all finite-dimensional distributions. Now we describe the behavior of a single component X_ξ in time based on the definitions (16) of U^T and 1.6 of $\widetilde{Y^\theta}$ and $\widetilde{\tau}$.

Theorem 1 (scaled component process) *Fix a label* $\xi \in \Xi$.

(a) **(convergence)** *There is a* $\{0,1\}$*–valued process* U^∞ *on the (macroscopic backward) time interval* $[0,1)$ *such that*

$$U^T \overset{\text{fdd}}{\Longrightarrow} U^\infty \quad as \quad T \to \infty.$$

(b) **(characterization of U^∞)** *For* $k \geq 0$ *and* $0 \leq \beta_0 < \beta_1 < \ldots < \beta_k < 1$:

$$\mathbb{P}\Big\{U^\infty_{\beta_0} = \ldots = U^\infty_{\beta_k} = 1\Big\} \;=\; E\Big(\widetilde{Y^\theta}(\beta_0)\cdots\widetilde{Y^\theta}(\beta_k)\Big) \;=\; E\big(\widetilde{Y^\theta}(\beta_k)\big)^{k+1}.$$

Consequently, the distribution of U^∞ *is a mixture of Bernoulli product laws: First realize the transformed Fisher-Wright diffusion* $\widetilde{Y^\theta}$ *and then build the product law with marginals*

$$\big(1 - \widetilde{Y^\theta}(\beta)\big)\,\delta_0 + \widetilde{Y^\theta}(\beta)\,\delta_1\,, \qquad 0 \leq \beta < 1. \tag{17}$$

In particular, the one-dimensional marginals $\mathcal{L}\big(U^\infty_\beta\big)$ *are given by* (11), *for all* $\beta \in [0,1)$.

(c) **(qualitative description of U^∞)** *Consider the holding time*

$$h_U \;:=\; \sup\big\{\beta \in [0,1);\, U^\infty_\beta = U^\infty_0\big\} \in (0,1)$$

of the initial state U^∞_0. *Then*

$$\mathcal{L}\big([U^\infty_0, h_U]\big) = \mathcal{L}\big([\widetilde{Y^\theta}(0), \widetilde{\tau}]\big).$$

Furthermore, beyond h_U *the process* U^∞ *is a "mixture" of non-stationary* 0-1*–noise: For* $\partial \in \{0,1\}$ *and the* $\beta_1, \ldots, \beta_k$ *as in* (b),

$$\begin{aligned}&\mathcal{L}\Big\{\Big[U^\infty_{\beta_1}, \ldots, U^\infty_{\beta_k}\Big]\;\Big|\; U^\infty_0 = \partial,\;\, h_U < \beta_1\Big\}\\ &= E\Big\{\textstyle\prod_{i=1}^k\Big[\big(1 - \widetilde{Y^\theta}(\beta_i)\big)\delta_0 + \widetilde{Y^\theta}(\beta_i)\delta_1\Big]\;\Big|\; \widetilde{Y^\theta}(0) = \partial,\;\, \widetilde{\tau} < \beta_1\Big\}.\end{aligned}$$

Remark 1.12 Note that the holding time h_U is *measurable* on the σ–algebra of all (backward) paths with a non-empty starting interval of constant value. ◇

The theorem says three things:
(i) If we look back from time N^T in time scale $N^{\beta T}$, the component we focus on has been "close" to its state ∂ for a time of random order β of magnitude.
(ii) This order is (strictly) positive and coincides in law with the holding time $\widetilde{\tau}$ of $\widetilde{Y^\theta}$.
(iii) Later changes occur in times of a smaller order of magnitude (conditional noise), within the correlation length.

Remark 1.13 (time average of components) Since the correlation length (in time) is small compared with the system's age, one could prove that objects of the form $t^{-1}\int_0^t \mathrm{d}s\; X_\xi(s)$ converge in law to θ. This is characteristic for the case of drift parameters $\{c_k\}$ not decaying exponentially fast (the analog of the $d = 2$ case in lattice models). Compare [CG83]. ◇

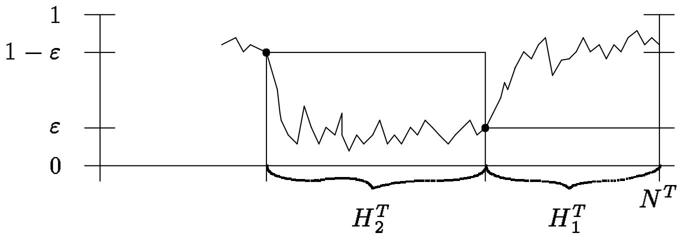

Figure 5: Alternating sequence of "holding times"

`An open problem` A very natural question is, how the holding times close to time points N^T behave in the limit $T \to \infty$. To be a bit more specific, for a fixed $\varepsilon \in (0, \frac{1}{2})$ we introduce a sequence of random (backward) times (see Figure 5):

$$\left\{N^T - \textstyle\sum_{1 \le i \le n} H_i^T;\ n \ge 1\right\}.$$

Here H_1^T is by definition the (first) hitting (backward) time of the boundary $[0, \varepsilon]$ if we start off at time N^T in $[1-\varepsilon, 1]$, or vice versa. H_2^T is then defined as the hitting (backward) time increment of the opposite boundary region starting at time $N^T - H_1^T$, etc. At this stage we agree to set a hitting time increment H_i^T (together with the subsequent H_j^T, $j > i$) equal to 0 if the time interval $[0, N^T]$ is exhausted.

For our purpose, the increments $H_1^T, H_2^T, \ldots$ may serve as the (backward) *holding times* of the component process X_ξ at the boundaries, since the fraction of time the component process spends in $[\varepsilon, 1-\varepsilon]$ converges to 0 in probability as $T \to \infty$; see Theorem 4 in [FG94].

Incorporating the scaling suggested by the result of Theorem 1, define the *rescaled holding times*

$$\widehat{H}_i^T := \frac{\log H_i^T}{T \log N} \tag{18}$$

(that is $N^{\widehat{H}_i^T T} = H_i^T$) which for our purpose describe the *order of magnitude* of H_i^T. What one would like to do now is the following:

- Show that $\mathcal{L}\{\widehat{H}_i^T;\, i \ge 1\}$ has a limiting law, say Γ.
- Identify the law Γ via the transformed Fisher-Wright tree.
- Show that Γ is concentrated on decreasing sequences.

In order to carry out such an analysis, which involves joint laws of holding times rescaled by functions of different order of magnitude, requires more than controlling moments of the time-space diagram. What is needed is a representation of the interacting system via particle systems in the sense of the work of Donelly and Kurtz [DK96]. Such analysis is outside the scope of the present paper.

1.4 Spatial ball averages in time dependence

We want to combine the previous set-up describing a single component during time with our results in [FG94] about the spatial structure at a fixed (late) time, and this way to obtain a better picture how the clusters evolve in time. We approach this phenomenon from two angles. Namely in the present subsection we consider spatial ball averages in their time dependence whereas in the next one we shall deal with thinned-out time-space fields.

For fixed $\xi \in \Xi$ and $\alpha \in [0,1)$ consider the following *spatial ball averages*

$$V_\beta^{\alpha,T} := \frac{1}{N^{[\alpha T]}} \sum_{\zeta:\, \|\zeta-\xi\| \le \alpha T} X_\zeta\big(N^T - N^{\beta T}\big) = X_{\xi,[\alpha T]}\big(N^T - N^{\beta T}\big), \tag{19}$$

$0 \le \beta < 1$, as processes in the *macroscopic backward time* $\beta \in [0,1)$ (here $[r]$ refers to the integer part of r). As $T \to \infty$, a limiting process $V^{\alpha,\infty}$ on $[0,1)$ will exist whose law depends on α. Since $N^{\beta T} = o(N^T)$ (for $\beta < 1$ fixed), we stay again within the correlation length, and the one-dimensional marginal distribution of $V^{\alpha,\infty}$ is again independent of β but is now given by the law $\widetilde{Q^\theta}_\alpha$ of the transformed Fisher-Wright diffusion $\widetilde{Y^\theta}$ of (9) at α; see [FG94, Theorem 2].

The next theorem deals with this time-scaled process of spatial ball averages. Recall that $\mu \in \mathcal{T}_\theta$ and $0<\theta<1$.

Theorem 2 (time-scaled spatial ball averages) *Fix* $0 \le \alpha < 1$.

(a) (convergence) *There exists a* $[0,1]$*–valued process* $\{V_\beta^{\alpha,\infty};\ 0 \le \beta < 1\}$ *with*

$$V^{\alpha,T} \overset{\text{fdd}}{\Longrightarrow} V^{\alpha,\infty} \quad \text{as} \quad T \to \infty. \tag{20}$$

(b) (characterization of $V^{\alpha,\infty}$) *Fix* $k, m_0, \dots, m_k \ge 0$ *and* $0 =: \beta_0 < \cdots < \beta_k < 1$. *Then*

$$\mathsf{E}\big(V_{\beta_0}^{\alpha,\infty}\big)^{m_0} \cdots \big(V_{\beta_k}^{\alpha,\infty}\big)^{m_k} = \mathbf{E}^\theta\Big[\Big(\widetilde{\mathbf{Y}^\theta}_0(\alpha)\Big)^{m_0+\dots+m_{J-1}} \prod_{J \le i \le k} \Big(\widetilde{\mathbf{Y}^\theta}_{\beta_i}(\alpha)\Big)^{m_i}\Big]$$

with $\widetilde{\mathbf{Y}^\theta}$ *the transformed Fisher-Wright tree of Definition* 1.9, *and*

$$J := \min\{i;\ \alpha \le \beta_i,\ 1 \le i \le k+1\}. \tag{21}$$

(c) (qualitative description of $V^{\alpha,\infty}$)

(c1) *The marginal laws* $\mathcal{L}\big(V_\beta^{\alpha,\infty}\big)$ *are given by* $\widetilde{Q^\theta}_\alpha$ *of Definition* 1.6(c), *for all* $\beta \in [0,1)$.

(c2) *Consider* $h_V := \sup\{\beta \in [0,1);\ V_\beta^{\alpha,\infty} = V_0^{\alpha,\infty}\}$, *the holding time of* $V^{\alpha,\infty}$ *(recall Remark* 1.12*). Then* $\mathcal{L}(h_V) = \mathcal{L}(\alpha \vee \widetilde{\tau})$ *(see Definition* 1.6(d)*).*

(c3) *Beyond* h_V *the finite-dimensional distributions of* $V^{\alpha,\infty}$ *are equal to*

$$\mathcal{L}\Big\{\Big[V_{\beta_1}^{\alpha,\infty},\dots,V_{\beta_k}^{\alpha,\infty}\Big]\Big| h_V < \beta_1\Big\} = \mathcal{L}\Big\{\Big[\widetilde{\mathbf{Y}^\theta}_{\beta_1}(\alpha),\dots,\widetilde{\mathbf{Y}^\theta}_{\beta_k}(\alpha)\Big]\Big| (\alpha \vee \widetilde{\tau}) < \beta_1\Big\}$$

(with the β_i *from* (b)*). Here the r.h.s. is the following mixture of product laws:*

$$\int R^{\theta,\alpha}_{\beta_1,\dots\beta_k}\big(\mathrm{d}[\theta_1,\dots,\theta_k]\big)\ \widetilde{Q^{\theta_1}}_{\alpha/\beta_1} \times \cdots \times \widetilde{Q^{\theta_k}}_{\alpha/\beta_k},$$

with $R^{\theta,\alpha}_{\beta_1,\dots\beta_k}$ *the conditional distribution of* $\Big[\widetilde{Y^\theta}(\beta_1),\dots,\widetilde{Y^\theta}(\beta_k)\Big]$ *given* $(\alpha \vee \widetilde{\tau}) < \beta_1$, *and with* $\widetilde{Q^{\theta_i}}_{\alpha/\beta_i}$ *as in* (10).

This theorem says that the spatial ball average has remained in its terminal value at least a time of order $N^{\alpha T}$. However, this holding time is larger than α if the whole α–ball is covered by a 0– or 1–cluster at the terminal time N^T (this event has positive probability), in which case (depending on the random size of that cluster) the empirical mean had been in the same state as at time N^T for a random time. The order of magnitude is $\alpha \vee \widetilde{\tau}$. Looking back further gives us then conditionally (given $\alpha \vee \widetilde{\tau}$) independent observations since the time grid is too large to detect earlier and hence small holding times. Theorem 2 (c) combined with the conjectures at p. 11 and a result in [FG94] suggests that a specific value in the order of magnitude of the holding time of a component (viewed backwards from a late time point) corresponds to the existence of a cluster at that late time which has a *corresponding* order of magnitude. Roughly speaking, on the used macroscopic scales, the spatial cluster size gives the holding time of a typical component in that cluster. This will be made precise in Theorem 3 below.

1.5 Time-space thinned-out systems

A second approach to investigate the history of a spatial cluster found at time N^T and to relate the order of spatial size of the cluster to the order of the holding time of a component, is the following. Choose a spatial network of points having distances αT. Consider a new field obtained by observing the system through time only at this network of observation points. Do this however only in a network of time points which also spread apart suitably as the system ages. We formalize this point of view as follows which will verbally be explained in Remark 1.15.

Definition 1.14 (thinning procedures)

(a) **(inverse level shift operators $\mathcal{S}_n^{-1}$ and spatially thinned-out systems $\mathcal{S}_n^{-1}x$)** For $n \geq 0$, $\xi \in \Xi$ and $x \in [0,1]^{\Xi}$, set

$$(\mathcal{S}_n^{-1}x)_\xi := x_{\mathcal{S}_n^{-1}\xi} \quad \text{with} \quad (\mathcal{S}_n^{-1}\xi)_j := \begin{cases} \xi_{n+j} & \text{if} \quad j > n, \\ 0 & \text{if} \quad 1 \leq j \leq n. \end{cases} \tag{22}$$

(b) **(space-time thinned-out systems)** Fix $k \geq 0$ and $1 > \beta_1 > \cdots > \beta_k \geq \alpha > 0 =: \beta_0$. Set $\underline{\beta} := [\beta_1, ..., \beta_k]$ and

$$W_{\xi,i}^{\underline{\beta},\alpha,T} := \big(\mathcal{S}_{[\alpha T]}^{-1}X\big)_\xi \Big[\Big(N^T - \sum_{1 \leq i' \leq i} N^{\beta_{i'}T}\Big)_+\Big], \tag{23}$$

$\xi \in \Xi,\ 0 \leq i \leq k,\ T > 1.$ ◇

Remark 1.15 $\mathcal{S}_n^{-1}$ shifts all coordinates (levels) of ξ by n steps, and fills in the newly created coordinates by 0. Hence, $\xi = 0$ is a fixed point, and if $\|\xi\| = m \neq 0$ then $\|\mathcal{S}_n^{-1}\xi\| = m + n$. In particular, $\mathcal{S}_n^{-1}$ increases non-zero distances of pairs of labels by n. Applied to a whole configuration $x \in [0,1]^{\Xi}$, we can view $\mathcal{S}_n^{-1}x$ as a *spatially thinned-out system* since each fixed pair of labels has distance n.

For the fixed scaling parameters $\underline{\beta} \geq \alpha$, we consider $[\xi, i] \in \Xi \times \{0, ..., k\}$ as new, macroscopic space-"time" variables of the random fields $W^{\underline{\beta},\alpha,T}$. As $T \to \infty$, these fields will have a $\{0,1\}^{\Xi \times \{0,...,k\}}$–valued limiting field denoted by $W^{\underline{\beta},\alpha,\infty}$. It describes the evolution of clusters both in time and space. ◇

Theorem 3 (time-rescaled thinned-out systems) *Fix scaling parameters $\underline{\beta} \geq \alpha$ as in Definition* 1.14 (b).

(a) **(convergence)** *There exists a $\{0,1\}^{\Xi \times \{0,...,k\}}$–valued random field $W^{\underline{\beta},\alpha,T}$ on* $\Xi \times \{0, ..., k\}$ *such that*

$$W^{\underline{\beta},\alpha,T} \overset{\text{fdd}}{\Longrightarrow} W^{\underline{\beta},\alpha,\infty} \quad \textit{as} \quad T \to \infty.$$

(b) **(characterization of $W^{\underline{\beta},\alpha,\infty}$)** *Fix natural numbers $m_0, ..., m_k \geq 0$ and, for each i in $\{0, ..., k\}$, distinct labels $\xi_{i,1}, ..., \xi_{i,m_i}$ in Ξ. Then*

$$\mathbb{P}\Big(W^{\underline{\beta},\alpha,\infty}_{\xi_{i,j},i} = 1;\ 0 \le i \le k,\ 1 \le j \le m_i\Big) = \mathbf{E}^\theta \prod_{i=0}^k \big(\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha)\big)^{m_i} \tag{24}$$

with $\widetilde{\mathbf{Y}}^\theta$ the transformed Fisher-Wright tree of Definition 1.9 (and $\beta_0 = 0$).

(c) **(qualitative description of $W^{\underline{\beta},\alpha,\infty}$)** $\{W^{\underline{\beta},\alpha,\infty}_{\xi,i};\ \xi \in \Xi,\ 0 \le i \le k\}$ *is an associated collection of $\{0,1\}$–valued random variables. To describe its distribution, let $F^{\underline{\beta},\alpha}$ denote the law of the random vector $\{\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha);\ 0 \le i \le k\}$, and write $\underline{\partial}$ for the configuration identically equal to ∂. Then*

$$\begin{aligned}\mathcal{L}(W^{\underline{\beta},\alpha,\infty}) &= \int F^{\underline{\beta},\alpha}(\mathrm{d}[u_0, ..., u_k])\Big[\prod_{i=0}^k (1-u_i)\delta_0 + u_i\delta_1\Big]^\Xi \\ &= \sum_{\partial=0,1} P\Big(Y^\theta(\tau) = \partial,\ \tau \le \log(1/\beta_1)\Big)\delta_{\underline{\partial}} \\ &\quad + \mathbf{E}^\theta\Big\{\prod_{i=0}^k\Big[\big(1-\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha)\big)\delta_0 + \widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha)\delta_1\Big]^\Xi;\ \widetilde{\tau} < \beta_1\Big\}.\end{aligned}$$

Consequently, $W^{\underline{\beta},\alpha,\infty}$ is a "mixture" of independent fields; with probability $P(\widetilde{\tau} \ge \beta_1)$ it is even a constant field $\underline{\partial}$ (with random ∂).

Theorems 1 (c) and 3 (c) reflect the fact that clusters have a space-time extension with an order of magnitude (α, α) where α is random. That is, the spatial cluster size is αT (in the hierarchical distance), whereas a "typical" component of that cluster lived for a time $N^{\alpha T}$. Or turned around, at time N^T, spatial clusters of size αT have an age of order $N^{\alpha T}$. Hence, in the time-space diagram of the process viewed back from the end N^T in an exponential time scale, we see at large times clusters of a size comparable with a square of a random size.

Remark 1.16 Both marginals of the fields are mixtures of product laws, and the mixing distributions are expressed via Fisher-Wright tree quantities. ◇

The most important feature of our analysis is that the large scale behavior of our model does not depend on the diffusion coefficient g, and in particular the transformed Fisher-Wright tree is an *universal object* in the class of models considered:

Corollary 1.17 (universality) *The limiting objects U, V, and (in the sense of finite-dimensional distributions) W depend essentially on the initial density $\theta \in (0,1)$, but are otherwise independent of the "input parameters" $a > 0$, $g \in \mathcal{G}^0$ and $\mu \in \mathcal{T}_\theta$ of the interacting diffusion X, and of the parameter N of the label set Ξ.*

1.6 Strategy of proofs and outline

The proofs of the Theorems 1–3 will follow the strategy to first reduce the general results by coupling and comparison techniques to the case of interacting Fisher-Wright diffusions starting in a product law. Then we can use a time-space duality relation with a delayed coalescing random walk ϑ with (deterministic) immigration, their approximation by an (instantaneous) coalescing random walk η with immigration, and scaling limits for the latter model.

For this purpose, in Section 2 we study some random walk systems, in particular coalescing random walks. In Section 3 we introduce an extension of Kingman's coalescent. We call this object $\widetilde{\Lambda}$ an ensemble of log-coalescents. In Section 4 it occurs in certain scaling limits of coalescing random walks (e.g., Theorem 4 at p. 27). On the other hand, it is in duality with the transformed Fisher-Wright tree $\widetilde{\mathbf{Y}}^{\theta}$ (Theorem 5 in Section 4, p. 31), which is our crucial object for the description of the space-time structure of interacting diffusions. In Section 5 other basic techniques like the duality of X and ϑ, coupling and moment comparison are compiled, culminating in the universal conclusion Theorem 6 at p. 37. In Section 6 we finally prove our Theorems 1–3 and with Theorem 7 (p. 39) a rather general version of a scaling limit for thinned-out X–systems.

2 Preliminaries: On coalescing random walks

A basic tool for our study of the interacting Fisher-Wright diffusion X will be a time-space duality relation with a delayed coalescing random walk with immigration. As a preparation for this, in the present section we develop the relevant random walk models and some of their properties.

2.1 Coalescing random walk with immigration

`Random walk` Z `on the hierarchical group` Ξ Let $Z = \{Z_t;\, t \geq 0\}$ denote the continuous-time (right-continuous) *random walk* in Ξ with *jump rate*

$$\kappa := \frac{aN^2}{N^2 - 1} \tag{25}$$

(where a is the drift parameter $a \equiv c_k$ of the interacting diffusion of Definition 1.1 and N the "degree of freedom" in the hierarchical group Ξ) and *jump probabilities*

$$p_{\xi,\zeta} := \frac{1}{N^{2\|\zeta-\xi\|}}\,, \quad \xi \neq \zeta, \quad \text{hence} \quad p_{\xi,\xi} \equiv \tfrac{N-1}{N}\,. \tag{26}$$

Let Z^{ξ} refer to Z starting with $Z(0) = \xi \in \Xi$ (at time 0). The law of $Z = Z^{\xi}$ is denoted by P^{ξ}. For convenience sometimes we also write $Z(t)$ instead of Z_t (similarly we proceed for other processes).

We recall from [FG94, Lemma 2.21 and Proposition 2.37] that Z is a *recurrent* random walk and that the *hitting time* distribution of the origin starting from a fixed point $\xi \neq 0$ has tails of order $1/\log t$ as $t \to \infty$. For a detailed study of this random walk we refer to Section 2 of [FG94].

`Delayed coalescing random walk` ϑ Let $\vartheta = \{\vartheta_{\xi}(t);\, \xi \in \Xi,\, t \geq 0\}$ denote the (right-continuous) *delayed coalescing random walk* in Ξ with *coalescing rate* $b > 0$ (which corresponds to the diffusion parameter of the interacting Fisher-Wright diffusion, recall (5)). By definition, in the delayed coalescing random walk ϑ the particles move according to independent random walks of the previous subsection *except* when two particles meet. In the case of such a collision, as long as the two particles are at the same site, they attempt to coalesce to a single particle with (exponential) rate b.

Write ϑ^{ψ} if ϑ starts (at time 0) with $\psi \in \Psi$. Here $\Psi \subset \mathbb{Z}_+^{\Xi}$ denotes the set of all those particle configurations $\psi = \{\psi_{\xi}\,;\, \xi \in \Xi\}$ which are finite: $\|\psi\| := \sum_{\xi} \psi_{\xi} < \infty$. The configurations ψ with $\|\psi\| = 1$ (unit configurations) are denoted by δ^{ξ} where $\xi \in \Xi$ is the position of the particle. Set

$$\operatorname{supp}\psi := \bigl\{\xi \in \Xi;\ \psi_{\xi} > 0\bigr\}. \tag{27}$$

For a detailed description and discussion of ϑ we refer to § 3.a in [FG94] where the model is called coalescing random walk with delay. (ϑ is the dual of the interacting Fisher-Wright diffusion, see (64) at p. 34 below.)

`Coalescing random walk` η Write $\eta = \eta^{\varphi},\ \varphi \in \Phi$, for the (instantaneous) *coalescing random walk* obtained by formally setting the coalescing rate b to ∞. Here Φ denotes the set of all (finite) populations $\varphi \in \Psi$ with at most one particle at each site, that is $\varphi_\xi \leq 1$ for all ξ; see § 3.c in [FG94] for a detailed exposition. (Recall that η is the dual of the voter model on Ξ with interaction described by $\kappa\, p_{\xi,\zeta}$ of (25) and (26); see Liggett [Lig85, Chapter 5].)

By an abuse of notation (no confusion will be possible), the distributions of η^{φ} and ϑ^{ψ} are written as $\mathbf{P}^{\varphi}$ and $\mathbf{P}^{\psi}$, respectively.

`Delayed coalescing random walk with immigration` As introduced above, the delayed random walk ϑ^{ψ} starts at time $t_0 = 0$ with $\vartheta(0) = \psi$. Now we modify the model in the following way. Consider a finite sequence $t_0, ..., t_k \in \mathbb{R}$ of (deterministic) time points and related (deterministic) populations $\psi^0, ..., \psi^k \in \Psi$, respectively. Start the delayed random walk at time $t^* := t_0 \wedge ... \wedge t_k$ with the related population ψ^*, but in addition let the related populations ψ^i *immigrate* at the remaining time points $t_i \neq t^*,\ i = 0, ..., k$. The resulting (right-continuous) *delayed coalescing random walk with (deterministic) immigration* is again denoted by ϑ but we exhibit the immigration parameters in the notation as follows:

$$\vartheta = \vartheta^{\psi^0,...,\psi^k}_{t_0,...,t_k}, \quad \mathbf{P}^{\psi^0,...,\psi^k}_{t_0,...,t_k}, \qquad t_0, ..., t_k \in \mathbb{R}, \quad \psi^0, ..., \psi^k \in \Psi.$$

In particular, the starting time point is also viewed as an immigration time point. Of course, in the case $k = 0$ and $t_0 = 0$ we are back to the original delayed coalescing random walk: $\mathbf{P}^{\psi}_0 = \mathbf{P}^{\psi}$.

Note that this family of (time-inhomogeneous) Markov processes has an obvious *generalized time-homogeneity property:*

$$\mathbf{P}^{\psi^0,...,\psi^k}_{t_0,...,t_k}\Big\{\vartheta_{t_k+t} \in \cdot \;\Big|\; \vartheta_{t_k-} = \psi'\Big\} = \mathbf{P}^{\psi'+\psi^k}\{\vartheta_t \in \cdot\}, \tag{28}$$

$t_0, ..., t_{k-1} \leq t_k,\ \ t \geq 0,\ \psi^0, ..., \psi^k, \psi' \in \Psi$.

`Coalescing random walk with immigration` Similarly we define η, the (instantaneous) *coalescing random walk with immigration* (where $b = \infty$) and use the notation

$$\eta = \eta^{\varphi^0,...,\varphi^k}_{t_0,...,t_k}, \quad \mathbf{P}^{\varphi^0,...,\varphi^k}_{t_0,...,t_k}, \qquad t_0, ..., t_k \in \mathbb{R}, \quad \varphi^0, ..., \varphi^k \in \Phi.$$

These processes have a generalized time-homogeneity property analogous to (28). In this case one should have in mind a picture as shown in Figure 6.

The delayed coalescing random walk process with immigration is in a time-space duality with the interacting Fisher-Wright diffusion process, see Proposition 5.1 at p. 34, whereas the (instantaneous) coalescing random walk with immigration is in a time-space duality with the voter model on Ξ. The word time-space refers here to the fact that we consider the whole path up to time t. (In the case of $\Xi = \mathbb{Z}^d$ with interaction determined by the simple random walk kernel $p_{\xi,\zeta}$, the latter time-space duality was developed in Cox and Griffeath [CG83] using the name "frozen" random walks instead of ones with "immigration".)

2.2 Basic coupling

Throughout the paper it will be useful to define the relevant random walk models on a *common* probability space. For comparison we shall also need a system $Z^{\chi^0,...,\chi^k}_{t_0,...,t_k}$ of *independent* random

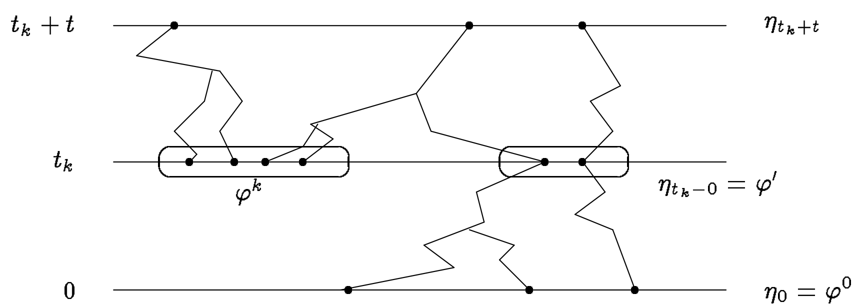

Figure 6: Coalescing random walk with immigration ($0 = t_0 < t_k < t_k + t,\ k = 1$)

walks with immigrating populations $\chi^0, ..., \chi^k \in \Psi$ at the times $t_0, ..., t_k$, respectively, defined as Ψ-valued process in the obvious way. Finally we give the following *basic coupling principle:*

Construction 2.1 (basic coupling) *Choose a basic probability space $[\Omega, \mathcal{F}, \mathcal{P}]$ in such a way that it supports all three (time-inhomogeneous) Markov families*

$$Z^{\chi^0,...,\chi^k}_{t_0,...,t_k}, \quad \vartheta^{\psi^0,...,\psi^k}_{t_0,...,t_k}, \quad \text{and} \quad \eta^{\varphi^0,...,\varphi^k}_{t_0,...,t_k},$$

where $k \geq 0$, $t_0, ..., t_k \in \mathbb{R}$, $\chi^0, ..., \chi^k, \psi^0, ..., \psi^k \in \Psi$ and $\varphi^0, ..., \varphi^k \in \Phi$, and that these families satisfy

$$Z^{\chi^0,...,\chi^k}_{t_0,...,t_k}(t) \ \geq\ \vartheta^{\psi^0,...,\psi^k}_{t_0,...,t_k}(t) \ \geq\ \eta^{\varphi^0,...,\varphi^k}_{t_0,...,t_k}(t), \qquad t \geq t^* := t_0 \wedge ... \wedge t_k\,,$$

whenever $\chi^i \geq \psi^i \geq \varphi^i$ for all $0 \leq i \leq k$.

Proof (existence of the basic coupling) First construct a probability space which supports the family of *independent* random walks with immigration $Z^{\chi^0,...,\chi^k}_{t_0,...,t_k}$. Then at time t^* we start $\|\chi^*\|$ independent walks placed according to the related χ^*, at all the remaining times t_i we additionally start $\|\chi^i\|$ independent walks placed according to χ^i. But in addition every immigrating particle (including at time t^*) gets an internal degree of freedom, by definition one of the numbers $0, 1$ or 2. The rules are as follows: If the immigrating particle belongs to one of the φ^i it gets the 0-mark, in the case of particles from $\psi^i - \varphi^i$ we adjoin the mark 1, and for $\chi^i - \psi^i$ we take 2. The mark of a particle is preserved during its evolution except for the following two situations:

- If two particles meet which have both the mark 0, then one of them (chosen at random) instantaneously gets the mark 1.
- If a pair of particles with mark in $\{0, 1\}$ (except if both are 0) stays at the same site, then at exponential rate b one of them having mark 1 is chosen at random (if we have two of them) and increases it's mark from 1 to 2. Here we let all possible pairs (at the same site) act independently.

Then at time $t \geq t^*$ count the particles as follows:

$$\begin{aligned}
\widehat{Z}^{\chi^0,...,\chi^k}_{t_0,...,t_k}(t) \quad &:= \quad \text{particles of all marks}\\
\widehat{\vartheta}^{\psi^0,...,\psi^k}_{t_0,...,t_k}(t) \quad &:= \quad \text{particles with marks 0 or 1}\\
\widehat{\eta}^{\varphi^0,...,\varphi^k}_{t_0,...,t_k}(t) \quad &:= \quad \text{particles with mark 0.}
\end{aligned}$$

Apparently these processes satisfy $\widehat{Z}(t) \geq \widehat{\vartheta}(t) \geq \widehat{\eta}(t),\ \ t \geq t^*$, and are a version of Z, ϑ, η as wanted. □

Note that the trivariate process $[Z, \vartheta, \eta]$ is *not* Markov. (In defining $[\widehat{Z}, \widehat{\vartheta}, \widehat{\eta}]$ by deleting the internal marks, the Markov character is lost.)

2.3 Approximation by (instantaneously) coalescing walks

Doubtless, (instantaneous) coalescing random walks with immigration are easier to handle than the corresponding delayed ones. On the other hand, we want to show now that in our context of a recurrent Z asymptotically the delayed coalescing random walk with immigration can be replaced without loss of generality by the corresponding system with instantaneous coalescence, and we will widely use this later on. (This implies in particular that the clustering properties of interacting Fisher-Wright diffusions are independent of the diffusion parameter $b > 0$.)

On an *intuitive level* this equivalence is justified by the following argument: If two particles do not meet, then the coalescing rate b is irrelevant and can be set to ∞. On the other hand, once two particles meet and do not coalesce before one of them jumps away, then by *recurrence* they will meet again and again until they will finally coalesce. (Caution: This heuristic argument has to be refined since it does not take into account that one of these two particles could meanwhile be "absorbed" by another particle.)

To put this idea on a firm base, first associate with each $\psi \in \Psi$ the *"truncated"* element $\psi \wedge 1 \in \Phi$ defined by $(\psi \wedge 1)_\xi := \psi_\xi \wedge 1,\ \ \xi \in \Xi$. The following result is a refinement and generalization of the approximation Proposition 3.6 of [FG94].

Proposition 2.2 (approximation of ϑ by η) *Fix integers $m_0, ..., m_k \geq 1,\ k \geq 0$. For $t > 1$, consider populations*

$$\psi^i = \psi^i(t) \in \Psi \quad \textit{with} \quad \left\|\psi^i\right\| = m_i\,, \qquad 1 \leq i \leq k,$$

and time points

$$s_0(t) < \cdots < s_k(t) < s_{k+1}(t) \quad \textit{with} \quad s_j(t) - s_i(t) \underset{t\to\infty}{\longrightarrow} \infty \quad \textit{if} \quad j > i.$$

Then on our basic probability space $[\Omega, \mathcal{F}, \mathcal{P}]$ (recall Construction 2.1), the event

$$\vartheta^{\psi^0,\dots,\psi^k}_{s_0,\dots,s_k}(s_{k+1}) \;=\; \eta^{\psi^0\wedge 1,\dots,\psi^k\wedge 1}_{s_0,\dots,s_k}(s_{k+1}) \tag{29}$$

has $\mathcal{P}$-probability converging to 1 as $t \to \infty$. (Sometimes we do not display the t-dependence.)

Remark 2.3 The approximate equivalence of ϑ and η explains via duality, why (in the recurrent case) interacting Fisher-Wright diffusions and the voter model on Ξ have a similar large scale behavior. ◇

Proof The proof proceeds by induction over k, the number of immigration time points.

1° (*initial step of induction*) Let $k = 0$. Then the processes are time-homogeneous, and for simplicity we may set $s_0(t) \equiv 0$. We treat this case $k = 0$ by doing again an induction, namely over the number m_0 of initial particles. For convenience, write $m_0 =: m,\ \psi^0 =: \psi$. Without loss of generality we may assume that $s_1(t) = N^t$ is satisfied (otherwise change the notation of $\psi(t)$). Fix representations $\psi(t) =: \delta^{\zeta(1,t)} + \cdots + \delta^{\zeta(m,t)}$. Trivially, the claim holds for $m = 1$. For the induction step, recall the coupling Construction 2.1 and assume that the statement is true for some $m-1 \geq 1$. Write

$$E_m := \left\{ \vartheta^{\psi}(N^t) = \eta^{\psi \wedge 1}(N^t) \right\}$$

for the event (29) (in the case $k=0$). Define E^m_{m-1} as E_m except replacing ψ by $\psi - \delta^{\zeta(m,t)}$. For a fixed $i < m$, let $C_{i,t}(s)$ and $M_{i,t}(s)$ denote the events that the walks $Z_{\zeta(i,t)}$ and $Z_{\zeta(m,t)}$ *coalesce* respectively *meet* by time s.

Let $\sigma(t)$ denote the first collision time of $Z^{\zeta(i,t)}$ and $Z^{\zeta(m,t)}$ after at least one of them jumped away from its initial state. Recall that the difference of the independent walks $Z^{\zeta(i,t)}$ and $Z^{\zeta(m,t)}$ is a random walk of the same kind except twice the jump rate. Define $\alpha(t)$ by $\alpha(t)t = \|\zeta(i,t) - \zeta(m,t)\|$. By the hitting probability Proposition 2.43 of [FG94], we have for $\gamma \in (0,1)$ fixed,

$$\mathcal{P}\left\{ N^t - N^{\gamma t} \leq \sigma(t) < N^t \right\} \xrightarrow[t \to \infty]{} 0. \tag{30}$$

(In fact, apply this proposition twice, namely with $\beta(t) \equiv 1$ and $\varrho(t) \equiv -\infty$ or $\varrho(t) \equiv \gamma$, respectively.)

Consider a subsequence $t' \to \infty$ such that the limit $\alpha(\infty) := \lim_{t' \to \infty} \alpha(t')$ exists in $[0, \infty]$. If $\alpha(\infty) \geq 1$ then by the same proposition we have

$$\mathcal{P}\left\{ \sigma(t') < N^{t'} \right\} \xrightarrow[t' \to \infty]{} 0. \tag{31}$$

In the opposite case $\alpha(\infty) < 1$, the latter probability has a positive limit, and (30) implies

$$\mathcal{P}\left\{ M_{i,t'}(N^{t'} - N^{\gamma t'}) \,\middle|\, M_{i,t'}(N^{t'}) \right\} \xrightarrow[t' \to \infty]{} 1.$$

But then due to *recurrence* of the random walk (cf. Lemma 2.21 in [FG94]) we conclude

$$\mathcal{P}\left\{ C_{i,t'}(N^{t'}) \,\middle|\, M_{i,t'}(N^{t'}) \right\} \xrightarrow[t' \to \infty]{} 1$$

and therefore

$$\mathcal{P}\left\{ M_{i,t'}(N^{t'}) \setminus C_{i,t'}(N^{t'}) \right\} \xrightarrow[t' \to \infty]{} 0. \tag{32}$$

On the other hand, from (31) we know that (32) holds also under $\alpha(\infty) \geq 1$. Summarizing (32) is true whenever $t = t' \to \infty$.

Dropping in notation the time argument N^t, we use the decomposition

$$E^m_{m-1} = \left(E^m_{m-1} \cap \bigcup_{i<m} M_{i,t} \right) \cup \left(E^m_{m-1} \cap \mathcal{C} \bigcup_{i<m} M_{i,t} \right) \tag{33}$$

(where $\mathcal{C}A$ denotes the complement of the event A). By the induction hypothesis, $\mathcal{P}\{E^m_{m-1}\}$ tends to 1 as $t \to \infty$, so the probability of the event on the r.h.s. of (33) tends to 1. By (32) we can replace in that event $\bigcup_{i<m} M_{i,t}$ by $\bigcup_{i<m} C_{i,t}$ to get still

$$\mathcal{P}\left\{ \left(E^m_{m-1} \cap \bigcup_{i<m} C_{i,t} \right) \cup \left(E^m_{m-1} \cap \mathcal{C} \bigcup_{i<m} M_{i,t} \right) \right\} \xrightarrow[t \to \infty]{} 1.$$

This finishes the proof by induction on m since the latter event implies E_m. Consequently the claim in the proposition holds in the case $k=0$.

2° (*induction step*) Using that the pair $[\vartheta, \eta]$ is a simple functional of a (bivariate) Markov process (see § 2.2) and exploiting generalized time-homogeneity as in (28), the induction step is similar to the argument for $k=0$ by considering the process starting with the configuration at the moment of the k-th immigration. □

2.4 Speed of spread of random walks

Our random walk Z in Ξ has the following property: At time scale N^t the speed of growth of the norm $\|Z(N^t)\|$ of $Z(N^t)$ is of order 1 as $t \to \infty$. To formulate with Lemma 2.6 below a more precise statement, for $r, c \geq 0$, set

$$\ell(r) := \frac{N \vee \log[r]}{\log N}, \tag{34}$$

and introduce the subsets

$$\overline{\Xi}[r,c] := \Big\{\xi \in \Xi;\ \|\xi\| \leq [r] + c\,\ell(r)\Big\}, \quad \Xi[r,c] := \Big\{\xi \in \Xi;\ \Big|\|\xi\| - [r]\Big| \leq c\,\ell(r)\Big\}, \tag{35}$$

of Ξ which consists of all labels ξ, up to a specific logarithmic error, of at most or exactly norm $[r]$. Note that the ring $\Xi[r,c]$ is contained in the ball $\overline{\Xi}[r,c]$, which is non-decreasing in r, and that both are non-decreasing in c. These sets have the following simple property.

Lemma 2.4 (spread of sums) *Fix constants $\alpha, \beta, c, d \geq 0$ with $\alpha < \beta$. For $t > 1$ let $\xi(t)$ in $\overline{\Xi}[\alpha t, c]$ and $\zeta(t)$ in $\Xi[\beta t, d]$ be given. Then, for all t sufficiently large,*

$$\xi(t) + \zeta(t) \in \Xi[\beta t, d]. \tag{36}$$

Remark 2.5 (cancellation) The assumption $\alpha < \beta$ cannot be dropped. For instance, if $\alpha = \beta = 1$ and $c = d = 0$ as well as $\xi(t) := -\zeta(t)$ then $\xi(t) + \zeta(t) \equiv 0 \notin \Xi[t,0]$. $\diamond$

Proof From the definition (35) we conclude

$$\|\xi(t)\| \leq [\alpha t] + c\,\ell(\alpha t), \quad \|\zeta(t)\| \geq [\beta t] - d\,\ell(\beta t).$$

Then $\alpha < \beta$ yields $\|\xi(t)\| < \|\zeta(t)\|$ for all $t \geq t_0$ say. Hence, $\|\xi(t) + \zeta(t)\| = \|\zeta(t)\|$ for these t by the definition of addition in Ξ. Consequently, (36) holds for $t \geq t_0$. $\square$

The announced speed property of our random walk now reads as follows. Note that we choose the initial state $\xi(t)$ of the random walk $Z^{\xi(t)}$ itself t-dependent.

Lemma 2.6 (walk speed) *Fix non-negative constants α, α' and positive constants β, ε, c. For $t > 1$, let $\varrho(t) \in \big[-\infty, \beta - \frac{\varepsilon}{t}\big]$, $\xi(t) \in \overline{\Xi}[\alpha t, c]$, and $\zeta(t) \in \overline{\Xi}[\alpha' t, c]$ be given. In the case $\alpha > \beta$, require even that $\xi(t)$ and $\xi(t) + \zeta(t)$ both belong to $\Xi[\alpha t, c]$. Then*

$$P\Big\{\zeta(t) + Z^{\xi(t)}\big(N^{\beta t} - N^{\varrho(t) t}\big) \in \Xi\big[(\alpha \vee \alpha' \vee \beta)t\,,\ 2c\big]\Big\} \underset{t\to\infty}{\longrightarrow} 1. \tag{37}$$

In particular, if $\|Z(0)\|$ is t-dependent and has a speed of order α then the speed of $\|Z^{\xi(t)}(N^{\beta t})\|$ is of order $\alpha \vee \beta$ as $t \to \infty$; that is, the time correction term $N^{\varrho(t)t}$ is negligible.

Proof Without loss of generality, in (37) we may set $\zeta(t) \equiv 0$. In fact, $\zeta(t) + Z^{\xi(t)}$ coincides in law with $Z^{\xi(t)+\zeta(t)}$, and $\xi(t) \in \overline{\Xi}[\alpha t, c]$, as well as $\zeta(t) \in \overline{\Xi}[\alpha' t, c]$ imply $\xi(t) + \zeta(t) \in \overline{\Xi}[(\alpha \vee \alpha')t, c]$, so in the case $\alpha \leq \beta$ we can rename $\xi(t), \zeta(t)$ and α. Moreover, in the case $\alpha > \beta$ we additionally assumed $\xi(t), \xi(t) + \zeta(t) \in \Xi[\alpha t, c]$, so again it is justified to rename $\xi(t)$ and $\zeta(t)$.

Now, under $\zeta(t) \equiv 0$, the case $\alpha \leq \beta$ directly follows from Lemma 2.26 in [FG94] (with ζ, s, r replaced by $\xi(t), \beta t, \varrho(t)t$, respectively). It remains to treat $\alpha > \beta$. For the moment, consider the walk Z^0 starting at the origin 0 of Ξ. By the proved part of the lemma, we may assume that $Z^0\big(N^{\beta t}-N^{\varrho(t)t}\big)$ belongs to $\Xi[\beta t, 2c]$. Then, by Lemma 2.4, for t sufficiently large, $Z^0\big(N^{\beta t}-N^{\varrho(t)t}\big)+\xi(t)\in\Xi[\alpha t,\, c]$. Hence, $Z^{\xi(t)}\big(N^{\beta t}-N^{\varrho(t)t}\big)\in\Xi[\alpha t,\, c]\subseteq\Xi[\alpha t,\, 2c]$ with probability converging to 1 as $t\to\infty$. This finishes the proof. □

Remark 2.7 (non-cancellation) In the case $\alpha\le\beta$, the cancellation effect of Remark 2.5 cannot happen in the situation of Lemma 2.6, since there is negligible probability that the walk will meet a prescribed point at a particular late time. ◇

2.5 Speed of spread of coalescing random walks

The above speed property of families of single random walks (Lemma 2.6) has consequences for the coalescing random walk with immigration, since we are interested in the latter system at late times and for time-dependent initial and immigrating populations. To describe the situation we need some notation (which is verbally explained below):

Definition 2.8 (spreading multi-colonies) Fix integers $\ell, m_0\,,\ldots,m_\ell\ge 0$ and non-negative constants $\alpha_0\,,\ldots,\alpha_\ell, c$. Write $\underline{\alpha}:=[\alpha_0\,,\ldots\alpha_\ell]$ and $\underline{m}:=[m_0\,,\ldots,m_\ell]$. For $t>1$, denote by $\overline{\Phi}_t[\,\underline{\alpha}\,,\underline{m}\,;c]$ the set of all those populations $\varphi=\varphi(t)\in\Phi$ which can be represented as $\varphi=\varphi^0+\cdots+\varphi^\ell$ where the $\varphi^j=\varphi^j(t)\in\Phi$, $0\le j\le\ell$, have the following properties (recall (35)):

(a) $\|\varphi^j(t)\|\equiv m_j$.

(b) If $\delta^\xi\le\varphi^j$ then $\xi=\xi(t)$ has to belong to $\overline{\Xi}[\alpha_j t\,,c]$.

(c) If $\delta^\xi+\delta^\zeta\le\varphi^j$ then we must have $\xi-\zeta\in\Xi[\alpha_j t\,,c]$.

(d) If $\delta^\xi\le\varphi^j$ and $\delta^\zeta\le\varphi^{j'}$ where $j\ne j\prime$ then $\xi-\zeta\in\Xi\big[(\alpha_j\vee\alpha_{j\prime})t\,,c\big]$.

If in (b) the balls $\overline{\Xi}[\alpha_j t\,,c]$ are replaced by the smaller rings $\Xi[\alpha_j t\,,c]$ then write $\Phi_t[\,\underline{\alpha}\,,\underline{m}\,;c]$ instead of $\overline{\Phi}_t[\,\underline{\alpha}\,,\underline{m}\,;c]$. (Note that $\Phi_t[\,\underline{\alpha}\,,\underline{m}\,;c]\subseteq\overline{\Phi}_t[\,\underline{\alpha}\,,\underline{m}\,;c]$.) Finally, write $\overline{\Phi}_t[\,\underline{\alpha}\,,\underline{\le m}\,;c]$ and $\Phi_t[\,\underline{\alpha}\,,\underline{\le m}\,;c]$ if in (a) only $\|\varphi^j(t)\|\equiv n_j\le m_j$ for some n_j . ◇

Consequently, a population $\varphi\in\Phi_t[\,\underline{\alpha}\,,\underline{m}\,;c]$ $\big($or $\varphi\in\overline{\Phi}_t[\,\underline{\alpha}\,,\underline{m}\,;c]\big)$, for which we often prefer to use the term *"multi-colony"*, is a superposition of $\ell+1$ subpopulations $\varphi^0,\ldots,\varphi^\ell$ of size $m_0\,,\ldots,m_\ell\ge 0$, respectively, with the following properties (up to specific logarithmic errors):

- Particles from the j-th subpopulation spread at (respectively at most at) speed α_j (see (b)).
- Pairs of particles from the j-th subpopulation spread with relative velocity α_j (cf. (c)).
- Mixed pairs of particles from $[\varphi^j,\varphi^{j'}]$ spread at relative speed $\alpha_j\vee\alpha_{j\prime}$ (see (d)).

Now we are in a position to formulate the main result of this subsection concerning the speed of spread of multi-colonies in the coalescing random walk with spreading immigrating populations. In simplified words it says the following: Suppose at times $s_i(t):=N^{\beta_i t}$, $i\le k$, we have an immigration by populations being a superposition consisting of ℓ_i+1 subpopulations of $m_{i,0}\,,\ldots,m_{i,\ell_i}$ particles with velocities determined by $\alpha_{i,0}\,,\ldots\alpha_{i,\ell_i}$, respectively. Then the terminal population at normalized time β_{k+1} is a superposition of subpopulations which spread apart with the velocities $\alpha_{i,j}\vee\beta_{k+1}$, $0\le j\le\ell_j$, (all except some logarithmic error terms and as described in Definition 2.8).

Proposition 2.9 (speed of spread for multi-colonies) *Fix integers k, $\ell_0, ..., \ell_k \geq 0$, constants $c \geq 1$, $0 \leq \beta_0 < \cdots < \beta_{k+1}$, a vector $\underline{\alpha}_i := [\alpha_{i,0}, ..., \alpha_{i,\ell_i}] \geq 0$ and an integer-valued vector $\underline{m}_i := [m_{i,0}, ..., m_{i,\ell_i}] \geq 0$. Assume that*

$$\alpha_{i',j} \neq (\alpha_{i,0} \vee \beta_{i'}), ..., (\alpha_{i,\ell_i} \vee \beta_{i'}) \qquad \text{if} \qquad 0 \leq i < i' \leq k, \quad 0 \leq j \leq \ell_{i'}. \tag{38}$$

Consider immigrating populations $\varphi^i(t)$ satisfying (recall Definition 2.8)

$$\varphi^i = \varphi^i(t) \in \overline{\Phi}_t[\underline{\alpha}_i, \underline{m}_i; c], \qquad t > 1, \quad 0 \leq i \leq k.$$

If for some i, $0 \leq i \leq k$, not all $\alpha_{i,0}, ..., \alpha_{i,\ell_i}$ are smaller than β_{i+1}, and, in the case $i > 0$, smaller than all of the $(\alpha_{i',0} \vee \beta_i), ..., (\alpha_{i',\ell_{i'}} \vee \beta_i)$, $0 \leq i' < i$, we even require $\varphi^i(t) \in \Phi_t[\underline{\alpha}_i, \underline{m}_i; c]$. Set $s_i = s_i(t) := N^{\beta_i t}$, $0 \leq i \leq k+1$. Then as $t \to \infty$, the event

$$\eta_{s_0, ..., s_k}^{\varphi^0, ..., \varphi^k}(s_{k+1}) \in \Phi_t\Big[\underline{\underline{\alpha}}^k \vee \beta_{k+1}, \leq\underline{\underline{m}}^k; 2^{k+1}c\Big] \tag{39}$$

has probability converging to 1. Here we abbreviated $\underline{\underline{m}}^k := [\underline{m}_0, ..., \underline{m}_k]$ and $\underline{\underline{\alpha}}^k \vee \beta_{k+1} := [\underline{\alpha}_0, ..., \underline{\alpha}_k] \vee \beta_{k+1}$.

Proof The proof will be by induction over k, the number of immigration time points.

$1°$ (*initial step of induction*) Consider $k = 0$ (no additional immigration), and drop the index 0 in notation. Consider a pair $\xi(t), \zeta(t)$ of "particles" taken from the initial population $\varphi = \varphi(t)$, that is $\delta^\xi + \delta^\zeta \leq \varphi$. Recall that the difference $\overline{Z} := Z^\xi - Z^\zeta$ of independent walks is a random walk in Ξ of the same kind but with twice the jump rate.

Now there are two cases possible: The pair ξ, ζ of particles originates

(i) from a subpopulation φ^j of φ related to the speed α_j,

(ii) from two different subpopulations φ^j and $\varphi^{j'}$ of φ (i.e. a "mixed" pair).

(i) By assumption on φ^j we have $\xi - \zeta \in \Xi[\alpha_j t, c]$ (recall condition (c) of Definition 2.8). Hence we may apply the walk speed Lemma 2.6 (with $\varrho = \beta_0$) to conclude that the event

$$\overline{Z}(s_1(t)) = Z^\xi(s_1(t)) - Z^\zeta(s_1(t)) \in \Xi[(\alpha_j \vee \beta_1)t, 2c] \tag{40}$$

has a probability converging to 1 as $t \to \infty$, and hence conditioning on this event is harmless. If now the coalescing mechanism is additionally applied (recall the coupling principle 2.1), then *on the event* (40) there are two cases. If the walks meet, then they coalesce, and we may apply the walk speed Lemma 2.6 to the surviving random walk starting with a particle ξ from φ^j which case has to be considered anyway (to check the condition (b) of Definition 2.8). Then we get the desired position $Z^\xi(s_1) \in \Xi[(\alpha_j \vee \beta_1)t, 2c]$. On the other hand, if the walks do not meet, then the pair ξ, ζ of particles survives by time s_1, and its relative position is in $\Xi[(\alpha_j \vee \beta_1)t, 2c]$, since we are in the event (40). Summarizing, the walks starting in the pair ξ, ζ from φ^j, end up at time $s_1(t)$ in a subpopulation corresponding to the (relative and absolute) speed $\alpha_j \vee \beta_1$.

(ii) Now consider a mixed pair ξ, ζ from $\varphi^j, \varphi^{j'}$. By assumption (recall condition (d) of Definition 2.8), it has relative speed $\alpha_j \vee \alpha_{j'}$, say α_j without loss of generality. Again by the walk speed Lemma 2.6, we may assume that (40) holds. Hence, we may continue to argue as in (i).

Combining (i) and (ii), we see that

$$\mathcal{P}\Big\{\eta_{s_0}^{\varphi^0}(s_1) \in \Phi_t[\underline{\underline{\alpha}}^0 \vee \beta_1, \leq\underline{\underline{m}}^0; 2c]\Big\} \xrightarrow[t \to \infty]{} 1.$$

2° (*induction step*) Consider $k \geq 1$. By the Markov property of the process η and generalized time-homogeneity as formulated in (28) at p. 16 for the process ϑ, the population from (39) can be thought of as arising from a process which starts in the population

$$\eta_{s_0,\dots,s_{k-1}}^{\varphi^0,\dots,\varphi^{k-1}}(s_k-) + \varphi^k =: \chi^k + \varphi^k \tag{41}$$

and running as a coalescing random walk for the time $N^{\beta_{k+1}t} - N^{\beta_k t}$.

Now we use that the claim is true for some $k-1 \geq 0$ (*induction hypothesis*). Then by (39) we may restrict our consideration to the case that χ^k belongs to

$$\Phi_t\Big[\underline{\underline{\alpha}}^{k-1} \vee \beta_k\,,\ \leq\underline{\underline{m}}^{k-1};\ 2^k c\Big]. \tag{42}$$

We take a pair ξ, ζ of particles from $\chi^k + \varphi^k$. The cases that both particles belong either to χ^k or to φ^k can be dealt with as in the first step of induction. The only difference is that we apply now the walk speed Lemma 2.6 with $\varrho = \beta_k$ instead of $\varrho = \beta_0$.

Thus it remains to consider the mixed case if one of the particles belongs to each of the sub-multi-populations. Say ξ belongs to χ^k whereas ζ is related to φ^k. Then $\xi \in \Xi[(\alpha_{i,j} \vee \beta_k)t\,,\ 2^k c]$ for some $i = 0, \dots, k-1$ and $j = 0, \dots, \ell_i$, and $\zeta \in \overline{\Xi}[\alpha_{k,j'}t\,,\ c]$ for some $j' = 0, \dots, \ell_k$. Now the condition (38) comes into the play, namely for $i' = k$. It guarantees that by the spread of sums Lemma 2.4 the speed of $\|\xi - \zeta\|$ can be determined by $\xi - \zeta \in \Xi\big[\big((\alpha_{i,j} \vee \beta_k) \vee \alpha_{k,j'}\big)t\,,\ 2^k c\big]$. Then one can continue as in the other two cases just described.

Summarizing, under the induction hypothesis, at the normalized time β_{k+1} we end up in the event as written in (39), with probability converging to one. This completes the proof by induction. $\square$

Remark 2.10 The condition $\varphi^i(t) \in \Phi_t[\underline{\alpha}_i\,, \underline{m}_i\,;\ c]$ says roughly that *all* absolute positions are of specified orders. This was required as soon as just one "violation" of parameter restrictions occurs. This is stronger than actually needed. But otherwise one would need a refined notation in order to describe the situation. $\diamond$

3 Ensemble of log-coalescents with immigration

In this section we study coalescing random walks with immigrating multi-colonies: We consider later and later time points and let the initial and immigrating populations spread apart. There exists a limiting object which we call an *ensemble of log-coalescents with immigration.* The crucial result is Theorem 4 at p. 27.

3.1 A log-coalescent $\widetilde{\lambda}$ with immigration

The purpose of this subsection is to introduce a death process on a logarithmic time scale, which we call the log-coalescent. In the next subsection we shall relate it with a scaling limit of a system of coalescing random walks with spreading initial populations (Proposition 3.2).

Start by recalling *Kingman's* [Kin82] *coalescent* $\lambda := \{\lambda(t);\ t \geq t_0\}$ *with coalescing rate* $b > 0$. By definition, this is a (time-homogeneous right-continuous Markov) death process starting at time $t_0 \in \mathbb{R}$ where a jump from $m \geq 0$ to $m-1$ occurs with rate $b\binom{m}{2}$. The process λ describes the evolution of finite populations of particles without locations, where each pair of particles coalesces into one particle with rate b, independently of all the other present pairs.

We agree to mean in the case $t_0 = -\infty$, that the process started with a (finite) state $\lambda(-\infty) \geq 0$ is defined as $\lambda(t) \equiv \lambda(-\infty) \wedge 1$ on $\mathbb{R}$.

From now on in this section we set the coalescing rate b to one (*standard* Kingman's coalescent). Next we define the *log-coalescent* $\widetilde{\lambda} = \{\widetilde{\lambda}(\alpha);\ \alpha \geq \alpha_0\}$ by setting

$$\widetilde{\lambda}(\alpha) := \lambda(\log \alpha), \qquad \alpha \geq \alpha_0 \geq 0. \tag{43}$$

This is a time-inhomogeneous Markov jump process starting at time α_0. (We call it the *log-coalescent*, to avoid confusion with Kingman's coalescent.)

The transition probabilities of $\widetilde{\lambda}$ are denoted by

$$p_\alpha^m(\beta, n) := \mathcal{P}\{\,\widetilde{\lambda}(\beta) = n \,|\, \widetilde{\lambda}(\alpha) = m\}, \qquad 0 \leq \alpha \leq \beta, \quad m, n \geq 0.$$

From the time-homogeneity of λ follows that

$$p_{c\alpha}^m(c\beta, n) \equiv p_\alpha^m(\beta, n), \qquad c > 0. \tag{44}$$

Since the transition probabilities of Kingman's coalescent λ can be calculated explicitly (see for instance Tavaré [Tav84, formula (6.1)]), we get for the transition probabilities of $\widetilde{\lambda}$ (restricting to $m \geq n \geq 1$):

$$p_\alpha^m(\beta, n) = \sum_{i=n}^{m} \frac{(-1)^{i-n}\,(2i-1)\,(i+n-2)!\,\binom{m}{i}}{n!\,(n-1)!\,(i-n)!\,\binom{m+i-1}{i}} \left(\frac{\alpha}{\beta}\right)^{\binom{i}{2}}, \quad \text{if} \quad 0 < \alpha \leq \beta, \tag{45}$$

and $\quad p_0^m(\beta, 1) \equiv 1 \quad$ if $\quad 0 = \alpha < \beta$.

In addition, we now allow a (deterministic) immigration of particles in the log-coalescent $\widetilde{\lambda}$.

Definition 3.1 (log-coalescent $\widetilde{\lambda}$ with immigration) At times $\alpha_0, ..., \alpha_\ell$ we let $m_0, ..., m_\ell$ particles immigrate, where the initial time point $\alpha_0 \wedge \cdots \wedge \alpha_\ell =: \alpha^*$ is again considered as an immigration time point. We write this *log-coalescent with immigration* and its transition probabilities as

$$\widetilde{\lambda}_{\underline{\alpha}}^{\underline{m}}(\beta) = \widetilde{\lambda}_{\alpha_0, ..., \alpha_\ell}^{m_0, ..., m_\ell}(\beta), \qquad p_{\underline{\alpha}}^{\underline{m}}(\beta, n) = p_{\alpha_0, ..., \alpha_\ell}^{m_0, ..., m_\ell}(\beta, n), \tag{46}$$

$\ell, n \geq 0, \quad \underline{\alpha} := [\alpha_0, ..., \alpha_\ell] \geq 0, \quad \beta \geq \alpha^*, \quad \underline{m} := [m_0, ..., m_\ell] \geq 0.$ ◇

Using the Markov property, one can easily establish the following *recursion formula:*

$$p_{\alpha_0, ..., \alpha_{\ell+1}}^{m_0, ..., m_{\ell+1}}(\beta, n) = \sum_{n'=1}^{m_0 + \cdots m_\ell} p_{\alpha_0, ..., \alpha_\ell}^{m_0, ..., m_\ell}(\alpha_{\ell+1}, n')\; p_{\alpha_{\ell+1}}^{n' + m_{\ell+1}}(\beta, n), \tag{47}$$

where $0 \leq \alpha_0, ..., \alpha_\ell \leq \alpha_{\ell+1} \leq \beta$, and where the last probability is given by (45) (process without immigration).

Obviously, (44) generalizes to

$$\widetilde{\lambda}_{c\underline{\alpha}}^{\underline{m}}(c\beta) \equiv \widetilde{\lambda}_{\underline{\alpha}}^{\underline{m}}(\beta), \quad p_{c\underline{\alpha}}^{\underline{m}}(c\beta, n) \equiv p_{\underline{\alpha}}^{\underline{m}}(\beta, n), \qquad c > 0. \tag{48}$$

3.2 Coalescing walk starting in spreading multi-colonies

Before we proceed further, in this subsection we demonstrate first in a simpler situation the role which is played by the log-coalescent with immigration. We restate a limit proposition concerning a coalescing random walk starting in (spreading) multi-colonies. In fact, Proposition 3.28 of [FG94] (which is analogous to Theorem 6 in [CG86]), with the now obvious identification of the limit probabilities, can be specialized as follows (formally we also include the case $m_i = 0$). Recall the rings $\Xi[r, c]$ of (35).

Proposition 3.2 (scaling limit for multi-colonies) *Fix non-negative integers $\ell, m_0, ..., m_\ell$, and $\varepsilon > 0$, $c \geq 1$, $0 \leq \alpha_0, ..., \alpha_\ell \leq \beta$ with $\beta > 0$. For $t > 1$ let $\varrho(t) \in \left[-\infty, \beta - \frac{\varepsilon}{t}\right)$. Moreover, for $0 \leq j \leq \ell$ let finite populations*

$$\varphi^j(t) = \delta^{\zeta^{j,1}(t)} + \cdots + \delta^{\zeta^{j,m_j}(t)} \in \Phi$$

be given with the property that

$$\zeta^{j,u}(t) - \zeta^{j',v}(t) \in \Xi\left[(\alpha_j \vee \alpha_{j'})t, c\right] \quad \textit{whenever} \quad [j,u] \neq [j',v], \tag{49}$$

and that the superposition $\varphi(t) := \varphi^0(t) + \cdots + \varphi^\ell(t)$ belongs to Φ. Then

$$\mathbf{P}^{\varphi(t)}\Big(\eta\big(N^{\beta t} - N^{\varrho(t)t}\big) = n\Big) \underset{t\to\infty}{\longrightarrow} p^{m_0, ..., m_\ell}_{\alpha_0, ..., \alpha_\ell}(\beta, n), \qquad n \geq 0,$$

with p the transition probability of the log-coalescent with immigration, satisfying the recursion formula (47).

Roughly speaking, start the coalescing random walk η with a superposition of $\ell+1$ subpopulations $\varphi^0, ..., \varphi^\ell$ where pairs of particles from $\varphi^j(t)$ spread with the relative velocity α_j whereas pairs from different subpopulations φ^j and $\varphi^{j'}$ spread with the relative speed $\alpha_j \vee \alpha_{j'}$. Then the number of particles at the late time $N^{\beta t}$ is approximately given by the log-coalescent $\widetilde{\lambda} = \widetilde{\lambda}^{m_0, ..., m_\ell}_{\alpha_0, ..., \alpha_\ell}$ at time β, with immigration of $m_0, ..., m_\ell$ particles at times $\alpha_0, ..., \alpha_\ell$, respectively. Note that only requirements on the *relative* position of particles in the initial populations are involved (in contrast to the scaling limit Theorem 4 below on the coalescing random walk with *immigrating* multi-colonies).

Remark 3.3 If the condition $\alpha_0, ..., \alpha_\ell \leq \beta$ in Proposition 3.2 is violated by some α_j then the walks starting with particles of this speed α_j cannot react by time $N^{\beta t}$ (with probability converging to 1 as $t \to \infty$). So they simply evolve independently, and in the limit these particles have to be added to the number of particles arising from the log-coalescent. Consequently, that condition is natural in that it is adapted to the actual *range of interaction* of the coalescing random walk. ◇

3.3 Ensembles $\widetilde{\Lambda}$ of log-coalescents with immigration

In this subsection we introduce the limiting object for coalescing random walks with immigration of spreading populations (multi-colonies). To avoid repeatedly cumbersome notation, we formulate a condition which we call the $\alpha{\leq}\beta$–Condition (recall Remark 3.3).

Condition 3.4 ($\alpha{\leq}\beta$–condition) Fix integers $k, \ell_0, ..., \ell_k \geq 0$, constants $0 \leq \beta_0 < \cdots < \beta_{k+1}$, vectors $\underline{\alpha}_i := [\alpha_{i,0}, ..., \alpha_{i,\ell_i}] \geq 0$ and $\underline{m}_i := [m_{i,0}, ..., m_{i,\ell_i}] \geq 0$, and suppose

$$\underline{\alpha}_0 \leq \beta_1 \qquad \text{and} \qquad \underline{\alpha}_i \leq \beta_i, \quad 1 \leq i \leq k,$$

(reading as $\alpha_{0,0} \leq \beta_1$ etc.) ◇

We now want to introduce what we call an *ensemble of log-coalescents with immigration*, see Figure 7. It will be used in the next subsection to describe a more general version of Proposition 3.2 above, namely a scaling limit for the coalescing random walk with *immigrating* multi-colonies.

Roughly speaking several log-coalescents with immigration evolve independently until they reach certain prescribed *deterministic* times $\beta_1 < \cdots < \beta_k$, respectively. In addition, we have a *tagged population* (related to the horizontal lines in the figure). From the times $\beta_1, ..., \beta_k$ on, the respective log-coalescents start to interact with the tagged population. (Recall that the coalescing rate b was set to 1.)

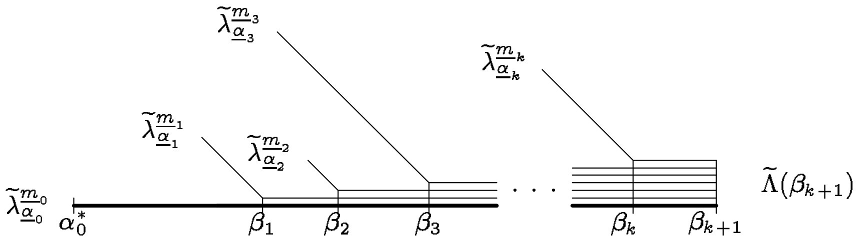

Figure 7: Ensemble $\widetilde{\Lambda}$ of log-coalescents with immigration

Definition 3.5 (ensemble $\widetilde{\Lambda}$ of log-coalescents with immigration)

(a) **(parameters)** Fix a constant $c \geq 1$. Suppose the $\alpha{\leq}\beta$–Condition 3.4. Set

$$\underline{\underline{\alpha}} = \underline{\underline{\alpha}}^k := [\underline{\alpha}_0, \dots, \underline{\alpha}_k], \quad \underline{\underline{m}} = \underline{\underline{m}}^k := [\underline{m}_0, \dots, \underline{m}_k], \quad \underline{\beta} = \underline{\beta}^k := [\beta_1, \dots, \beta_k] \tag{50}$$

and $\alpha_i^* := \alpha_{i,0} \wedge \cdots \wedge \alpha_{i,\ell_i}$.

(b) **(independent branches/random immigrants)** Let $\widetilde{\lambda}^{\underline{m}_1}_{\underline{\alpha}_1}, \dots, \widetilde{\lambda}^{\underline{m}_k}_{\underline{\alpha}_k}$ be *independent* log-coalescents with immigration, running during the time intervals $[\alpha_i^*, \beta_i]$, $1 \leq i \leq k$, respectively.

(c) **(tagged population)** We now define a process (*tagged population*) on the time interval $[\alpha_0^*, \beta_{k+1}]$ *given* the log-coalescents (branches) $\widetilde{\lambda}^{\underline{m}_1}_{\underline{\alpha}_1}, \dots, \widetilde{\lambda}^{\underline{m}_k}_{\underline{\alpha}_k}$ with immigration. On the subinterval $[\alpha_0^*, \beta_1)$ we set it equal to $\widetilde{\lambda}^{\underline{m}_0}_{\underline{\alpha}_0}$, that is we (only) run a log-coalescent with immigration determined by $\underline{m}_0, \underline{\alpha}_0$. Then at the time interval $[\beta_1, \beta_{k+1}]$ we continue with the log-coalescent, but with an additional *immigration* of $\widetilde{\lambda}^{\underline{m}_1}_{\underline{\alpha}_1}(\beta_1), \dots, \widetilde{\lambda}^{\underline{m}_k}_{\underline{\alpha}_k}(\beta_k)$ particles at the times $\beta_1, \dots, \beta_k$, respectively.

(d) **(ensemble $\widetilde{\Lambda}$ of log-coalescents with immigration)** Using the ingredients (a) – (c), we denote by

$$\widetilde{\Lambda}(\beta) = \widetilde{\Lambda}^{\underline{\underline{\alpha}};\underline{\underline{m}}}_{\underline{\beta}}(\beta), \qquad \alpha_0^* \leq \beta \leq \beta_{k+1},$$

the number of living particles in the tagged population (with random immigration) at time β. In particular, $\widetilde{\Lambda}(\beta_{k+1})$ denotes the *terminal number* of particles in the whole system. We call $\widetilde{\Lambda}$ the *ensemble of log-coalescents with immigration and parameters* $\underline{\underline{\alpha}}, \underline{\underline{m}}, \underline{\beta}$. ◇

By a generalized time-homogeneity, the following *recursion formula* holds:

$$\mathcal{P}\Big(\widetilde{\Lambda}^{\underline{\underline{\alpha}}^k;\underline{\underline{m}}^k}_{\underline{\beta}^k}(\beta_{k+1}) = n\Big) = \sum_{n'=0}^{\infty} \mathcal{P}\Big(\widetilde{\Lambda}^{\underline{\underline{\alpha}}^{k-1};\underline{\underline{m}}^{k-1}}_{\underline{\beta}^{k-1}}(\beta_k) = n'\Big)\, p^{\underline{m}_k, n'}_{\underline{\alpha}_k, \beta_k}(\beta_{k+1}, n), \tag{51}$$

$k \geq 1$, $n \geq 0$. Note that the number of non-vanishing terms in the sum is bounded by $\sum_{i,j} m_{i,j}$, hence finite. Clearly, (48) generalizes to the following *homogeneity property:*

$$\widetilde{\Lambda}^{c\underline{\underline{\alpha}};\underline{\underline{m}}}_{c\underline{\beta}}(c\beta) \equiv \widetilde{\Lambda}^{\underline{\underline{\alpha}};\underline{\underline{m}}}_{\underline{\beta}}(\beta), \qquad c > 0. \tag{52}$$

Definition 3.6 (ensemble of coalescents without immigration) If $\ell_0 = \cdots = \ell_k = 0$ in the $\alpha{\leq}\beta$–Condition 3.4 and in Definition 3.5 then we call $\widetilde{\Lambda}$ an ensemble of coalescents *without* immigration, and write simply $\widetilde{\Lambda}^{\underline{\alpha};\underline{m}}_{\underline{\beta}}$. ◇

Remark 3.7 Note that the term "*without* immigration" refers only to the fact that *within* the (randomly) immigrating branches of Definition 3.5 (b) no immigration occurs. ◇

3.4 Coalescing walk with immigration: Multi-colonies

Now we will formulate the announced scaling limit theorem for the coalescing random walk with immigrating multi-colonies spreading moderately (recall the Definition 2.8 at p. 21): On a macroscopic scale the latter behaves as an ensemble of log-coalescents with immigration. (Recall (50).)

Theorem 4 (scaling limit with immigrating multi-colonies) *Fix a constant $c \geq 1$, and suppose the $\alpha{\leq}\beta$–Condition 3.4. Consider immigrating multi-colonies*

$$\varphi^i = \varphi^i(t) \in \overline{\Phi}_t\big[\,\underline{\alpha}_i\,, \underline{m}_i\,;\,c\big], \qquad t > 1, \quad 0 \leq i \leq k.$$

Set $s_i = s_i(t) := N^{\beta_i t}$, $0 \leq i \leq k+1$. Then for the terminal population size of the related coalescing random walk with immigration we get

$$\mathcal{L}\left(\left\|\eta^{\varphi^0,\ldots,\varphi^k}_{s_0,\ldots,s_k}(s_{k+1})\right\|\right) \underset{t\to\infty}{\Longrightarrow} \mathcal{L}\Big(\widetilde{\Lambda}(\beta_{k+1})\Big) \tag{53}$$

with $\widetilde{\Lambda} = \widetilde{\Lambda}^{\underline{\underline{\alpha}};\underline{\underline{m}}}_{\underline{\beta}}$ the ensemble of log-coalescents with immigration.

Remark 3.8 Note that the limit process $\widetilde{\Lambda}$ is independent of the jump rate κ of the underlying random walk and the parameter N of Ξ. – Also, the limits are *non-degenerate* except some boundary cases as e.g. if $k = \ell_0 = 0$ and $m_{0,0} = 1$ implying $\widetilde{\Lambda}(\beta) \equiv 1$. – Recall that the limit law satisfies the recursion formula (51). ◇

Proof of Theorem 4 The proof is by induction over the number k of immigration time points. The case $k = 0$ (no immigration) follows from the scaling limit Proposition 3.2 for multi-colonies at p. 25 (with $\varrho = \beta_0$), since $\varphi^0(t) \in \overline{\Phi}_t[\,\underline{\alpha}_0\,, \underline{m}_0\,;\,c]$ is sufficient for the assumptions there.

For the induction step, with $k \geq 1$ consider

$$\mathbf{P}^{\varphi^0(t),\ldots,\varphi^k(t)}_{s_0(t),\ldots,s_k(t)}\Big\{\,\big\|\eta\big(s_{k+1}(t)\big)\big\| = n\Big\}, \qquad n \geq 0.$$

By the Markov property and generalized time-homogeneity as in (28) at p. 16 we can rewrite the expression as

$$= \mathbf{E}^{\varphi^0,\ldots,\varphi^{k-1}}_{s_0,\ldots,s_{k-1}}\, \mathbf{P}^{\eta(s_k-)+\varphi^k}\Big\{\big\|\eta'(s_{k+1}-s_k)\big\| = n\Big\} \tag{54}$$

with η' denoting an independent copy of η. According to the speed of spread of multi-colonies Proposition 2.9 at p. 22 we may assume that the subpopulation $\eta\big(s_k(t)-\big)$ belongs to the set

$$\Phi_t\Big[\beta_k\,;\,\leq\underline{\underline{m}}^{k-1};\,2^k c\Big]$$

whereas for the other subpopulation, $\varphi^k(t) \in \overline{\Phi}_t\big[\,\underline{\alpha}_k\,, \underline{m}_k\,;\,c\big] \subseteq \overline{\Phi}_t\big[\,\underline{\alpha}_k\,, \underline{m}_k\,;\,2^k c\big]$ holds by assumption. Moreover, by the walk speed Lemma 2.6 at p. 20, the relative speed of mixed pairs ξ, ζ of particles can uniformly be determined: $\xi - \zeta \in \Xi[\beta_k t\,,\,2^k c]$, since ξ arises from a walk starting at time s_{k-1} with a particle having a speed $\leq \beta_{k-1}$.

Altogether, the two subpopulations related to the two summands in $\eta(s_k-) + \varphi^k$ fulfill the requirements in the scaling limit Proposition 3.2 for multi-colonies (with $\varrho = \beta_k$). Hence, given $\|\eta\big(s_k(t)-\big)\| = n'$, the probability expression appearing in (54) has a limit which is given by $p^{\underline{m}_k, n'}_{\underline{\alpha}_k, \beta_k}(\beta_{k+1}, n)$ (recall (46) for the latter).

Now by the induction hypothesis the statement on the population sizes is true for some $k-1 \geq 0$. Then

$$\mathbf{P}^{\varphi^0,\dots,\varphi^{k-1}}_{s_0,\dots,s_{k-1}}\Big\{\big\|\eta(s_k-)\big\| = n'\Big\} \underset{t\to\infty}{\longrightarrow} \mathcal{P}\Big(\widetilde{\Lambda}^{\underline{\underline{\alpha}}^{k-1};\,\underline{\underline{m}}^{k-1}}_{\underline{\beta}^{k-1}}(\beta_k) = n'\Big).$$

Combined with the previous convergence statement for the probability conditioned on $\|\eta(s_k-)\| = n'$, we arrive at the r.h.s. of the recursion formula (51), since the number of terms over which we sum is finite. This completes the proof by induction. □

3.5 Coalescing walk with immigration: Colonies of common speed

Occasionally the $\alpha{\leq}\beta$–Condition 3.4 is not satisfied, therefore we prepare now a tool to treat such a situation. This comes up when at a sequence of time points *single* colonies immigrate which spread at a *common* speed α: On a macroscopic scale, by time α such coalescing random walk with immigration behaves like a system of non-interacting particles, and from time α on like an ensemble of log-coalescents without immigration (recall Definition 3.6).

Proposition 3.9 (immigrating colonies of common speed) *Fix integers $k, m_0, \dots, m_k \geq 0$, and constants $0 < \alpha < 1$, $0 =: \beta_0 < \cdots < \beta_{k+1} := 1$, $c \geq 1$. For $t > 1$ and $0 \leq i \leq k$ consider colonies $\varphi^i = \varphi^i(t) \in \Phi$ such that*

$$\|\varphi^i(t)\| \equiv m_i \quad \textit{and} \quad \varphi^0 + \cdots + \varphi^k \in \Phi_t\big[\alpha\,, m_0 + \cdots + m_k\,; c\big].$$

Put $s_i = s_i(t) := N^{\beta_i t}$, $0 \leq i \leq k$. Then

$$\mathcal{L}\bigg(\Big\|\eta^{\varphi^0,\dots,\varphi^k}_{s_0,\dots,s_k}(N^t)\Big\|\bigg) \underset{t\to\infty}{\Longrightarrow} \mathcal{L}\Big(\widetilde{\Lambda}(1)\Big)$$

with $\widetilde{\Lambda}$ the ensemble $\widetilde{\Lambda}^{[\alpha,\dots,\alpha]\,;\,[m_0+\cdots+m_{J-1}\,,\,m_J\,,\dots,m_k]}_{[\beta_J\,,\dots,\beta_k]}$ of log-coalescents without immigration, and with J defined in (21)*, p.* 12.

The limit object looks as follows: The tagged population and all the branches of $\widetilde{\Lambda}$ start at time α, namely with $m_0 + \cdots + m_{J-1}\,, m_J\,, \dots, m_k$ particles, respectively. They evolve independently as log-coalescents without immigration, until the branches coalesce with the tagged population at the times $\beta_J\,, \dots, \beta_k$, respectively.

Proof Since all the immigrating particles have absolute and relative speed α, by time $N^{\alpha t}$ none of them can interact by the walk speed Lemma 2.6. More precisely, by that lemma,

$$\eta^{\varphi^0(t),\dots,\varphi^k(t)}_{s_0(t),\dots,s_k(t)}(N^{\alpha t}) \in \Phi_t\big[\alpha, m_0 + \cdots + m_{J-1}\,, 2^J c\big]$$

with probability converging to 1 as $t \to \infty$. But starting with time $N^{\alpha t}$, we may apply Theorem 4, specialized to "single-colonies", to get the claim of the proposition. □

3.6 Coalescing walk with immigration: Exponential immigration time increments

Here we deal with a different time regime: Single populations with a common spreading speed immigrate, but now the immigration time *increments* are of the form $N^{\beta t}$, and actually of a *de*creasing order. The limit is again an ensemble of log-coalescents without immigration.

Proposition 3.10 (exponential immigration time increments) *Fix integers* $k, m_0, ..., m_k \geq 0$, *as well as constants*

$$1 > \beta_1 > \cdots > \beta_k \geq \alpha > 0 \tag{55}$$

and $c \geq 1$. *For* $t > 1$ *and* $0 \leq i \leq k$ *let colonies* $\varphi^i = \varphi^i(t) \in \overline{\Phi}_t[\alpha, m_i\,; c]$ *be given. Set*

$$s_i = s_i(t) := \textstyle\sum_{1 \leq i' \leq i} N^{\beta_{i'}\, t}, \qquad 0 \leq i \leq k.$$

Then

$$\mathcal{L}\Big(\Big\| \eta_{s_0,\ldots,s_k}^{\varphi^0,\ldots,\varphi^k}(N^t) \Big\| \Big) \underset{t\to\infty}{\Longrightarrow} \mathcal{L}\big(\widetilde{\Lambda}(1)\big) \tag{56}$$

with $\widetilde{\Lambda}$ *the ensemble* $\widetilde{\Lambda}_{[\beta_k,\ldots,\beta_1]}^{[\alpha,\ldots,\alpha]\,;\,[m_k,\ldots,m_0]}$ *of log-coalescents without immigration.*

In the limit object, the tagged population and all the branches of $\widetilde{\Lambda}$ start at time α, namely with $m_k, ..., m_0$ particles, respectively. They evolve independently as log-coalescents without immigration, until the branches coalesce with the tagged population at the times $\beta_k, ..., \beta_1$, respectively.

Proof The proof proceeds in two qualitatively different steps: First we analyze the evolution up to time $s_k(t)$, and then we provide the final step from time $s_k(t)$ to N^t.

$1°$(*initial population*) By the speed of spread Proposition 2.9 and the scaling limit Proposition 3.2, we conclude that after the first step:

$$\eta_{s_0}^{\varphi^0}(s_1-) \in \Phi_t[\beta_1, n_0\,; 2c] \quad \text{with random } n_0 = \widetilde{\lambda}_\alpha^{m_0}(\beta_1)$$

(with probability converging to 1 as $t \to \infty$). In the following time steps of macroscopic size $\beta_i < \beta_1$, this subpopulation $\eta_{s_0}^{\varphi^0}(s_1-)$ further behaves (asymptotically) as a system of *independent* random walks (walk speed Lemma 2.6), which at time s_k satisfies

$$\chi^0 = \chi^0(t) := \eta_{s_0}^{\varphi^0}(s_k-) \in \Phi_t[\beta_1, n_0\,; 2^k c]$$

(repeated use of Proposition 2.9).

$2°$(*second immigration*) By definition, $\varphi^1 \in \overline{\Phi}_t[\alpha, m_1\,; c]$ additionally immigrates at time s_1. By the parameter assumption (55), during the subsequent time increments, these new particles cannot interact with the subpopulation of $1°$ (Lemma 2.6). On the other hand, their own evolution is similar to that of the initial population: φ^1 results at time s_k into a subpopulation

$$\chi^1 \in \Phi_t[\beta_2, n_1\,; 2^{k-1}c] \subseteq \Phi_t[\beta_2, n_1\,; 2^k c] \quad \text{with random } n_1 = \widetilde{\lambda}_\alpha^{m_1}(\beta_2).$$

$3°$ (*all immigrants*) Continuing arguing in this way, at time s_k- we finally get k *independent* subpopulations

$$\chi^i \in \Phi_t[\beta_{i+1}, n_i\,; 2^k c] \quad \text{with random } n_i = \widetilde{\lambda}_\alpha^{m_i}(\beta_{i+1}), \qquad 0 \leq i \leq k-1,$$

(with probability converging to one).

$4°$ (*final step*) Define $\varrho(t)$ by $s_k(t) = N^{\varrho(t)t}$. For the final step from time s_k to N^t, we may apply the scaling limit Proposition 3.2 for multi-colonies with $\varphi^0, ..., \varphi^\ell$ replaced by $\chi^0, ..., \chi^{k-1}, \varphi^k$, given

$n_0, \dots, n_{k-1}$. In fact, also mixed pairs of particles from the total population at time s_k satisfy the spreading condition (49), by the walk speed Lemma 2.6. Therefore,

$$\mathcal{L}\Big(\eta^{\varphi^0,\dots,\varphi^k}_{s_0,\dots,s_k}(N^t)\Big) \underset{t\to\infty}{\Longrightarrow} \mathcal{L}\Big(\widetilde{\lambda}^{n_0,\dots,n_{k-1},m_k}_{\beta_1,\dots,\beta_k,\alpha}(1)\Big) = \mathcal{L}\Big(\widetilde{\lambda}^{m_k,n_{k-1},\dots,n_0}_{\alpha,\beta_k,\dots,\beta_1}(1)\Big)$$

where $[n_0, \dots, n_{k-1}]$ is random, is independent of the evolution, and equals in law with the independent vector

$$\Big[\widetilde{\lambda}^{m_0}_{\alpha}(\beta_1), \dots, \widetilde{\lambda}^{m_{k-1}}_{\alpha}(\beta_k)\Big].$$

But according to the Definition 3.5 of the ensemble of log-coalescents, specialized to the case without immigration, this limiting object can be described as claimed, finishing the proof. □

4 Duality of $\widetilde{\mathbf{Y}^\theta}$ and $\widetilde{\Lambda}$

In Theorem 4 (p. 27) of the previous section we learned that on a large space and time scale the *coalescing random walk with immigrating multi-colonies* can be described by an ensemble $\widetilde{\Lambda}$ of log-coalescents with immigration. In order to calculate probabilities for this limit process we use a duality relation with an object much simpler to handle. In fact, the main result of this section (Theorem 5) says that the limiting system is in duality with the *transformed Fisher-Wright tree* of Definition 1.9.

4.1 Duality of $\widetilde{\mathbf{Y}^\theta}$ and $\widetilde{\Lambda}$

Let $Y = \{Y(t);\ 0 \le t \le \infty\}$ denote the *Fisher-Wright diffusion with diffusion parameter* $b > 0$. By definition this is a diffusion process on the interval $[0,1]$ with generator determined by the differential operator $\frac{1}{2}\, b\,(r - r^2)\,\frac{\partial^2}{\partial r^2}$, $\quad 0 \le r \le 1$. Recall that the terminal state $Y(\infty) \in \{0,1\}$ is reached already after a finite time.

Consider the function $h(n,r) := r^n$, $n \ge 0$, $0 \le r \le 1$. If we apply the generator of Kingman's coalescent λ with coalescing rate $b > 0$, introduced in § 3.1, to $h(\cdot, r)$ then we get

$$b\binom{n}{2}\big[h(n-1,r) - h(n,r)\big] = \frac{1}{2}\, b\,(r - r^2)\,\frac{\partial^2}{\partial r^2}\, h(n,r). \tag{57}$$

Consequently, recalling the action of the Fisher-Wright generator on $h(n,\cdot)$, the generators of the (time-homogeneous) Markov processes λ and Y are in duality and we get the well-known *duality between Kingman's* [Kin82] *coalescent* λ *and the Fisher-Wright diffusion* Y (both with parameter b and starting at time 0):

$$E^n \theta^{\lambda(t)} = E^\theta Y^n(t), \qquad \theta \in [0,1], \quad n \ge 0, \quad t \ge 0, \tag{58}$$

(Tavaré [Tav84]).

Switch to the standard situation $b = 1$. Turning to a logarithmic scale, we will generalize this duality relation in Theorem 5 below. It will tell us that the generating function of the terminal number of particles in the ensemble $\widetilde{\Lambda}$ of coalescence with immigration can be expressed via moments of the transformed Fisher-Wright tree $\widetilde{\mathbf{Y}^\theta}$. (The definitions of $\widetilde{\mathbf{Y}^\theta}$ and $\widetilde{\Lambda}$ were given in 1.9 and 3.5 at pp. 7 and 26, respectively.)

Theorem 5 (duality of $\widetilde{\mathbf{Y}}^\theta$ and $\widetilde{\Lambda}$) *Suppose the $\alpha\le\beta$–Condition 3.4 at p. 25 with $\beta_0 := 0$ and $\beta_{k+1} := 1$. Set $\underline{\underline{m}} = \underline{\underline{m}}^k := [\underline{m}_0, ..., \underline{m}_k]$, $\underline{\underline{\alpha}} = \underline{\underline{\alpha}}^k := [\underline{\alpha}_0, ..., \underline{\alpha}_k]$, and $\underline{\beta} = \underline{\beta}^k := [\beta_1, ..., \beta_k]$. Then the generating function of the terminal number $\widetilde{\Lambda}(1)$ of particles in the ensemble $\widetilde{\Lambda} = \widetilde{\Lambda}_{\underline{\beta}^k}^{\underline{\underline{\alpha}}^k;\underline{\underline{m}}^k}$ of log-coalescents with immigration and parameters $\underline{\underline{\alpha}}, \underline{\underline{m}}, \underline{\beta}$ is given by*

$$\sum_{n=1}^{\infty} \mathcal{P}\Big(\widetilde{\Lambda}_{\underline{\beta}^k}^{\underline{\underline{\alpha}}^k;\underline{\underline{m}}^k}(1) = n\Big)\theta^n \;=\; \mathbf{E}^\theta\Big[\prod_{i=0}^{k}\prod_{j=0}^{\ell_i}\Big(\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha_{i,j})\Big)^{m_{i,j}}\Big], \tag{59}$$

$0 \le \theta \le 1$, with $\widetilde{\mathbf{Y}}^\theta$ the transformed Fisher-Wright tree of Definition 1.9.

Example 4.1 In the special case $\ell_i \equiv 0$, $m_{i,0} \equiv m_i \le 1$, the r.h.s. of (59) simplifies to

$$E\big(\widetilde{Y^\theta}(\beta_0)\big)^{m_0} \cdots \big(\widetilde{Y^\theta}(\beta_k)\big)^{m_k} \;=\; E\left(\widetilde{Y^\theta}(\beta_k)\right)^{m_0+\cdots+m_k} \tag{60}$$

with the transformed Fisher-Wright diffusion $\widetilde{Y^\theta}$ defined in (9). In fact, condition first on the trunk. Then all branches become conditionally independent. Next, for all i with $m_i = 1$, we can use the martingale property of the Fisher-Wright diffusion to replace the (conditional) expectation over the k independent branches by their termination points $\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\beta_i) = \widetilde{\mathbf{Y}}^\theta{}_0(\beta_i) = \widetilde{Y^\theta}(\beta_i)$, $1 \le i \le k$. This gives the l.h.s. of (60). Then apply again the martingale property. – Note in particular, that here the $\alpha_{i,0}$ are irrelevant. This is immediately clear from the ensemble of log-coalescents since there is only at most one particles in each branch, which cannot react before its termination time, hence its "age" is irrelevant. ◇

4.2 Proof of the duality Theorem 5

For convenience, as a preparation we first expose some elementary properties of the Fisher-Wright tree $\mathbf{Y}^\theta$ from Definition 1.7:

Lemma 4.2 (elementary properties of $\mathbf{Y}^\theta$) *With respect to $\mathbf{P}^\theta$:*

(a) **(exchangeability)** *Given a splitting point $\mathbf{Y}^\theta_\infty(s_i)$ for a branch, the corresponding branch $\mathbf{Y}^\theta_{s_i}$ and the trunk from s_i on have the same law:*

$$\Big[\mathbf{Y}^\theta_\infty(s_i),\, \big\{\mathbf{Y}^\theta_{s_i}(t);\, t \ge s_i\big\}\Big] \overset{L}{=} \Big[\mathbf{Y}^\theta_\infty(s_i),\, \big\{\mathbf{Y}^\theta_\infty(t);\, t \ge s_i\big\}\Big], \qquad k \ge i \ge 1.$$

(b) **(time-homogeneity)** *Fix $k \ge i > 1$. Given the σ-field $\mathcal{F}(s_i)$, the vector $\big[\mathbf{Y}^\theta_{s_{i-1}}, ..., \mathbf{Y}^\theta_{s_1}\big]$ of $i-1$ branches is equal in law to*

$$\Big[\big\{\mathbf{Y}^{\theta'}_{s_{i-1}-s_i}(t-s_i);\, t \ge s_{i-1}\big\}, ..., \big\{\mathbf{Y}^{\theta'}_{s_1-s_i}(t-s_i);\, t \ge s_1\big\}\Big]$$

where $\theta' := \mathbf{Y}^\theta_\infty(s_i)$.

(c) **(conditional independence)** *Fix $k > 1$. Then $\big[\mathbf{Y}^\theta_{s_k}, \big\{\mathbf{Y}^\theta_{s_{k-1}}, ..., \mathbf{Y}^\theta_{s_1}\big\}\big]$ is an independent pair, given $\mathcal{F}(s_k)$.*

For later reference, we rewrite Lemma 4.2 for the transformed Fisher-Wright tree $\widetilde{\mathbf{Y}}^\theta$ (introduced in Definition 1.9):

Lemma 4.3 (some elementary properties of $\widetilde{\mathbf{Y}}^\theta$) *With respect to* $\mathbf{P}^\theta$:

(a) (exchangeability) *For fixed* $i \in \{1, ..., k\}$,

$$\left[\widetilde{\mathbf{Y}}^\theta{}_0(\beta_i),\ \left\{\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha);\ \alpha \le \beta_i\right\}\right] \stackrel{L}{=} \left[\widetilde{\mathbf{Y}}^\theta{}_0(\beta_i),\ \left\{\widetilde{\mathbf{Y}}^\theta{}_0(\alpha);\ \alpha \le \beta_i\right\}\right].$$

(b) (homogeneity) *Fix* $1 < i \le k$. *Given* $\widetilde{\mathcal{F}}(\beta_i)$, *the vector of branches* $\left[\widetilde{\mathbf{Y}}^\theta{}_{\beta_1}, ..., \widetilde{\mathbf{Y}}^\theta{}_{\beta_{i-1}}\right]$ *is equal in law with*

$$\left[\left\{\widetilde{\mathbf{Y}}^{\theta'}{}_{\beta_1/\beta_i}(\alpha/\beta_i);\ \alpha \le \beta_1\right\}, ..., \left\{\widetilde{\mathbf{Y}}^{\theta'}{}_{\beta_{i-1}/\beta_i}(\alpha/\beta_i);\ \alpha \le \beta_{i-1}\right\}\right]$$

where $\theta' := \widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\beta_i) = \widetilde{\mathbf{Y}}^\theta{}_0(\beta_i) = \widetilde{Y}^\theta(\beta_i)$.

(c) (conditional independence) *Fix* $k > 1$. *Then* $\left[\left\{\widetilde{\mathbf{Y}}^\theta{}_{\beta_1}, ..., \widetilde{\mathbf{Y}}^\theta{}_{\beta_{k-1}}\right\}, \widetilde{\mathbf{Y}}^\theta{}_{\beta_k}\right]$ *is an independent pair, given* $\widetilde{\mathcal{F}}(\beta_k)$.

Proof of Theorem 5 We proceed again by induction on k, the number of immigration time points of $\widetilde{\Lambda}$ (the number of branches in $\widetilde{\mathbf{Y}}^\theta$).

1° (*initial step of induction*) In the case $k = 0$ the ensemble $\widetilde{\Lambda}$ reduces to a single log-coalescent $\widetilde{\lambda} = \widetilde{\lambda}^{\underline{m}^0}_{\underline{\alpha}_0}$ with immigration. By formula (6.2) in [FG94], the generating function related to its terminal number $\widetilde{\lambda}(1)$ is given by

$$\sum_{n=1}^{\infty} \mathcal{P}\left(\widetilde{\lambda}^{\underline{m}^0}_{\underline{\alpha}_0}(1) = n\right)\theta^n \ = \ \mathbf{E}^\theta \prod_{j=0}^{\ell_0} \left(\widetilde{Y}^\theta(\alpha_{0,j})\right)^{m_{0,j}}. \tag{61}$$

Recalling $\widetilde{\mathbf{Y}}^\theta{}_0 = \widetilde{Y}^\theta$ yields (59) in the case $k = 0$.

2° (*induction step*) Let $k \ge 1$. Then by the recurrence formula (51) and the homogeneity property (52) the l.h.s. of (59) can be written as

$$\sum_{n'} \mathcal{P}\left(\widetilde{\Lambda}^{\underline{\underline{\alpha}}^{k-1};\underline{\underline{m}}^{k-1}}_{\underline{\beta}^{k-1}}(\beta_k) = n'\right) \sum_n \mathcal{P}\left(\widetilde{\lambda}^{\underline{m}_k, n'}_{\underline{\alpha}_k, \beta_k}(1) = n\right)\theta^n. \tag{62}$$

By the initial step of induction (recall (61)), the innermost sum equals

$$\mathbf{E}^\theta\left[\left(\widetilde{\mathbf{Y}}^\theta{}_0(\beta_k)\right)^{n'} \prod_{j=0}^{\ell_k}\left(\widetilde{\mathbf{Y}}^\theta{}_0(\alpha_{k,j})\right)^{m_{k,j}}\right] = \mathbf{E}^\theta\left[\left(\widetilde{\mathbf{Y}}^\theta{}_0(\beta_k)\right)^{n'} \prod_{j=0}^{\ell_k}\left(\widetilde{\mathbf{Y}}^\theta{}_{\beta_k}(\alpha_{k,j})\right)^{m_{k,j}}\right],$$

where we used Lemma 4.3(a) (with $i = k$). Inserting this into (62), interchanging the expectation $\mathbf{E}^\theta$ with the summation, and further rearranging yields

$$\mathbf{E}^\theta\mathbf{E}^\theta\left\{\prod_{j=0}^{\ell_k}\left(\widetilde{\mathbf{Y}}^\theta{}_{\beta_k}(\alpha_{k,j})\right)^{m_{k,j}} \sum_{n'} \mathcal{P}\left(\widetilde{\Lambda}^{\underline{\underline{\alpha}}^{k-1}/\beta_k;\underline{\underline{m}}^{k-1}}_{\underline{\beta}^{k-1}/\beta_k}(1) = n'\right)\left(\widetilde{\mathbf{Y}}^\theta{}_0(\beta_k)\right)^{n'} \,\middle|\, \widetilde{\mathcal{F}}(\beta_k)\right\},$$

where we additionally used the homogeneity property (52) with $c = 1/\beta_k$.

Assume now that (59) is valid for some $k-1 \geq 0$ (*induction hypothesis*). Then the latter sum equals

$$\mathbf{E}^{\theta'}\left[\prod_{i=0}^{k-1}\prod_{j=0}^{\ell_i}\left(\widetilde{\mathbf{Y}^{\theta'}}_{\beta_i/\beta_k}(\alpha_{i,j}/\beta_k)\right)^{m_{i,j}}\right], \quad \text{where} \quad \theta' := \widetilde{\mathbf{Y}}^{\theta}{}_0(\beta_k).$$

By Lemma 4.3 (b) (with $i = k$), given $\widetilde{\mathcal{F}}(\beta_k)$, this coincides with

$$\mathbf{E}^{\theta}\left\{\prod_{i=0}^{k-1}\prod_{j=0}^{\ell_i}\left(\widetilde{\mathbf{Y}^{\theta}}_{\beta_i}(\alpha_{i,j})\right)^{m_{i,j}}\middle|\,\widetilde{\mathcal{F}}(\beta_k)\right\}.$$

Finally, by the conditional independence property 4.3 (c) we can write the resulting expression

$$\mathbf{E}^{\theta}\left[\mathbf{E}^{\theta}\left\{\prod_{j=0}^{\ell_k}\left(\widetilde{\mathbf{Y}^{\theta}}_{\beta_k}(\alpha_{k,j})\right)^{m_{k,j}}\middle|\,\widetilde{\mathcal{F}}(\beta_k)\right\}\mathbf{E}^{\theta}\left\{\prod_{i=0}^{k-1}\prod_{j=0}^{\ell_i}\left(\widetilde{\mathbf{Y}^{\theta}}_{\beta_i}(\alpha_{i,j})\right)^{m_{i,j}}\middle|\,\widetilde{\mathcal{F}}(\beta_k)\right\}\right]$$

as expected conditional expectation

$$\mathbf{E}^{\theta}\mathbf{E}^{\theta}\left\{\left[\prod_{j=0}^{\ell_k}\left(\widetilde{\mathbf{Y}^{\theta}}_{\beta_k}(\alpha_{k,j})\right)^{m_{k,j}}\right]\prod_{i=0}^{k-1}\prod_{j=0}^{\ell_i}\left(\widetilde{\mathbf{Y}^{\theta}}_{\beta_i}(\alpha_{i,j})\right)^{m_{i,j}}\middle|\,\widetilde{\mathcal{F}}(\beta_k)\right\}.$$

But this is equal to the r.h.s. of (59), finishing the proof by induction. □

5 Duality, Coupling and Comparison

In this section we compile some basic methods to prove limit theorems for the interacting diffusion X as introduced in § 1.1. The basic tools combined will allow us to prove the key result of this section, Theorem 6, which asserts the universality of the limits obtained for the special case of interacting Fisher-Wright diffusions starting with product initial laws. Furthermore, using Section 4 and 2 we actually see in Theorem 6 that everything boils down to coalescing random walks with immigration, an object studied in Section 3.

The methods needed are the following: a *time-space duality* of interacting Fisher-Wright diffusions which is in particular useful in the case of i.i.d. initial components, a *successful coupling* enabling us to generalize from product measure to any initial state in $\mathcal{T}_\theta$, and a *moment comparison* to provide the step from Fisher-Wright $g = bf,\ b > 0,$ to general diffusion coefficients g in $\mathcal{G}^0$.

5.1 Time-space duality of X and ϑ

It is convenient to write the defining equation (1) for X in the form

$$\mathrm{d}X_\xi(t) = \kappa\sum_{\zeta\in\Xi}\left(p_{\xi,\zeta} - \delta_{\xi,\zeta}\right)X_\zeta(t)\,\mathrm{d}t + \sqrt{g\left(X_\xi(t)\right)}\,\mathrm{d}w_\xi(t), \qquad \xi\in\Xi, \tag{63}$$

with migration rate κ and migration probabilities p defined in (25) and (26), respectively, and with $\delta_{\xi,\zeta} = 1$ if $\xi = \zeta$, and $\delta_{\xi,\zeta} = 0$ otherwise.

We now develop a time-space duality between the *interacting Fisher-Wright diffusion* X (with diffusion parameter $b > 0$) and the delayed coalescing random walk ϑ with immigration (with coalescing rate $b > 0$).

First recall that a single Fisher-Wright diffusion and Kingman's coalescent are in duality as written in (58). Taking into account that the drift term in the interacting diffusion (63) is related to a continuous time random walk determined by κq, relation (58) generalizes to *Shiga's* [Shi80] *duality relation* between the interacting Fisher-Wright diffusion X and the delayed coalescing random walk ϑ as follows (Ψ was defined before (27)):

$$\mathbb{E}_z^b X_t^\psi = \mathbf{E}^\psi z^{\vartheta(t)}, \qquad z \in [0,1]^\Xi, \quad \psi \in \Psi, \quad t \geq 0. \tag{64}$$

(Here the notation $z^\psi := \prod_{\xi \in \Xi} z_\xi^{\psi_\xi}$ is used.) This relates all the multivariate moments of X_t with the generating functions of $\vartheta(t)$.

Since we want to study not only the law of the interacting diffusion at a single time t, but rather the whole path up to time t, we actually need the distributions of the process X viewed backwards from a "late" time point, say t_{k+1}. Hence we want to calculate moments of the form $\mathbb{E}_z^b X_{t_{k+1}-t_0}^{\psi^0} \cdots X_{t_{k+1}-t_k}^{\psi^k}$ with backward time points $0 =: t_0 < t_1 < \ldots < t_{k+1}$ (viewed from t_{k+1}). These moments can again be expressed via generating functions of a delayed coalescing random walk but now with immigration of particles exactly at those fixed time points $t_1, \ldots, t_k$. Here is the needed generalization of duality to multiple time points (recall that $b > 0$):

Proposition 5.1 (time-space duality of X and ϑ) *For $z \in [0,1]^\Xi$, $k \geq 0$, $\psi^0, \ldots, \psi^k \in \Psi$ and $0 \leq t_0 < \ldots < t_{k+1}$ the following duality relation holds:*

$$\mathbb{E}_z^b X_{t_{k+1}-t_0}^{\psi^0} \cdots X_{t_{k+1}-t_k}^{\psi^k} = \mathbf{E}_{t_0,\ldots,t_k}^{\psi^0,\ldots,\psi^k} z^{\vartheta(t_{k+1})}. \tag{65}$$

Consequently, the duality formula (65) relates the moments of the interacting Fisher-Wright diffusion X (starting at z) of orders $\psi^0, \ldots, \psi^k$ at times looked *backwards* from t_{k+1}, namely at the times $t_{k+1} - t_0$, ..., $t_{k+1} - t_k$, with the generating functions of the delayed coalescing random walk ϑ with immigrating populations ψ^0, ..., ψ^k at the *forward* times t_0, ..., t_k, respectively.

Remark 5.2 Only the "antiton" order in the duality relation (65) is important, that is one can interchange the role of forward and backward times. $\diamond$

Proof of Proposition 5.1 The proof is by induction on k. For $k = 0$ we are back to the original duality relation (64) since X and ϑ without additional immigration are both time-homogeneous. Let $k \geq 1$. Apply the Markov property at the "earliest forward" time $t_{k+1} - t_k$, and the time homogeneity of X to get for the l.h.s. of (65)

$$\mathbb{E}_z^b X_{t_{k+1}-t_k}^{\psi^k} X_{t_{k+1}-t_0}^{\psi^0} \cdots X_{t_{k+1}-t_{k-1}}^{\psi^{k-1}} = \mathbb{E}_z^b X_{t_{k+1}-t_k}^{\psi^k} \mathbb{E}_{X(t_{k+1}-t_k)}^b \widehat{X}_{t_k-t_0}^{\psi^0} \cdots \widehat{X}_{t_k-t_{k-1}}^{\psi^{k-1}}$$

where $\widehat{X}$ is an independent copy of X. Now assume that the assertion (65) is true for some $k-1 \geq 0$ (instead of k). Applying (65) to $\widehat{X}$ we can continue with

$$= \mathbb{E}_z^b X_{t_{k+1}-t_k}^{\psi^k} \mathbf{E}_{t_0,\ldots,t_{k-1}}^{\psi^0,\ldots,\psi^{k-1}} X_{t_{k+1}-t_k}^{\vartheta(t_k)} = \mathbf{E}_{t_0,\ldots,t_{k-1}}^{\psi^0,\ldots,\psi^{k-1}} \mathbb{E}_z^b X_{t_{k+1}-t_k}^{\psi^k+\vartheta(t_k)}.$$

Apply the original duality relation (64) (that is the initial step of induction) to arrive at

$$= \mathbf{E}_{t_0,\ldots,t_{k-1}}^{\psi^0,\ldots,\psi^{k-1}} \mathbf{E}^{\psi^k+\vartheta(t_k)} z^{\vartheta'(t_{k+1}-t_k)}$$

where ϑ' is an independent copy of ϑ. The interior generating function expression can be reformulated using the generalized time-homogeneity as in (28). This finishes the proof of (65) by induction. $\square$

If one specializes (65) to a one-component space $\Xi = \{0\}$, then one gets the time-space duality relation between the Fisher-Wright diffusion and Kingman's coalescent with immigration. Such formulas occur already in the literature, see for instance Cox [Cox89, formula (6.5)].

5.2 Successful coupling in the Fisher-Wright case

Coupling will actually be used twofold. Namely in the first place to get rid of independence assumptions concerning the initial state $X(0)$ for which the duality (65) is still tractable. But also to truncate initial states in order to be able to handle some restricted interacting Fisher-Wright diffusions needed in § 5.3. To prepare for the second case we first want to modify a bit our basic model introduced in Definition 1.1.

Definition 5.3 (diffusion coefficients in $\mathcal{G}$) Let $\mathcal{G} \supset \mathcal{G}^0$ denote the set of all diffusion coefficients g which are defined as in $\mathcal{G}^0$ (recall Definition 1.1 (d) at p. 4) *except* that we require *strict* positivity of g on a non-empty *sub*interval of $(0, 1)$ *only*. Note that the definition of the interacting diffusion X as strong solution to (63) still makes sense for these general $g \in \mathcal{G}$. $\diamond$

Definition 5.4 (coupling principle) Fix $g \in \mathcal{G}$ and two initial laws μ, ν on $[0, 1]^{\Xi}$. Let Γ be a distribution on $[0, 1]^{\Xi} \times [0, 1]^{\Xi}$ with marginals μ, ν. Choose $[X(0), \widehat{X}(0)]$ according to Γ, and solve (63) separately starting with $X(0)$ and $\widehat{X}(0)$, respectively, but using the *same* collection $\{w_\xi \, ; \, \xi \in \Xi\}$ of driving standard Brownian motions (recall that we work with the unique *strong* solution of (63)). Then the bivariate process $[X, \widehat{X}]$ is called the *coupling of the interacting diffusions X and $\widehat{X}$ with diffusion coefficient g and joint initial law Γ*. Write $\mathbb{P}^g_\Gamma$ for its distribution, and $\mathbb{P}^g_{[x,y]}$ in the degenerate case $\Gamma = \delta_x \times \delta_y$. $\diamond$

We now use this coupling concept to control the effect of a particular *truncation of the initial state*. For this purpose, for $0 \le \varepsilon \le \frac{1}{3}$ and $z \in [0, 1]^{\Xi}$ define the *truncated configuration* $z^\varepsilon \in [\varepsilon, 1-\varepsilon]^{\Xi}$ by

$$z^{\varepsilon}_{\xi} := \varepsilon \vee z_\xi \wedge (1 - \varepsilon), \qquad \xi \in \Xi.$$

Moreover, if z is distributed according to μ then we write μ^ε for the *"truncated law"*, that is for the distribution of z^ε.

Lemma 5.5 (truncation of initial states) *Let* $0 \le \varepsilon \le \frac{1}{3}$.

(a) (error control) *For the coupling* $[X, \widehat{X}]$ *starting in* $[z, z^\varepsilon]$,

$$\mathbb{E}^g_{[z,z^\varepsilon]} \left| X_\xi(t) - \widehat{X}_\xi(t) \right| \le \varepsilon, \qquad g \in \mathcal{G}, \quad z \in [0, 1]^{\Xi}, \quad \xi \in \Xi, \quad t \ge 0.$$

(b) (truncation in $\mathcal{T}_\theta$) *If μ belongs to the set $\mathcal{T}_\theta$ of shift ergodic laws with intensity $\theta \in (0, 1)$, then the truncated μ^ε belongs to $\mathcal{T}_{\theta^\varepsilon}$ for some $\theta^\varepsilon \in [\varepsilon, 1 - \varepsilon]$ with $\theta^\varepsilon \to \theta$ as $\varepsilon \downarrow 0$.*

Proof For (a), see the proof of Lemma 4.6 in [FG94], whereas (b) is obvious. □

Now we come to the main point of this subsection concerning interacting Fisher-Wright diffusions, namely to recall Proposition 5.11 of [FG94]. It says, roughly speaking, that coupled processes started with the *same* initial density θ approach each other as time increases, due to the fact that the same driving Brownian motions are used:

Lemma 5.6 (successful coupling of interacting Fisher-Wright diffusions) *Assume that* $g = bf,\ b > 0$. *Let* $\mu, \nu \in \mathcal{T}_\theta$. *Then the coupling* $[X, \widehat{X}]$ *of interacting Fisher-Wright diffusions with joint initial law* $\mu \times \nu$ *is successful, that is*

$$\mathbb{E}^b_{\mu\times\nu}\big|X_0(t) - \widehat{X}_0(t)\big| \underset{t\to\infty}{\longrightarrow} 0.$$

Successful coupling will enable us to switch from product initial laws μ in $\mathcal{T}_\theta$ to general $\nu \in \mathcal{T}_\theta$.

5.3 Comparison with restricted Fisher-Wright diffusions

Since the limit processes U, V and W of the Theorems 1,2,3 do not depend on the diffusion coefficient $g \in \mathcal{G}^0$, our basic method to get this *universality* in g is a *comparison principle* with (restricted) interacting Fisher-Wright diffusions. This is a special case of a general comparison principle proved in Cox et al. [CFG96].

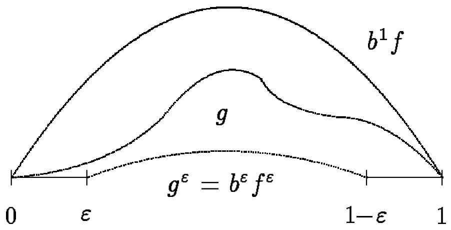

Figure 8: (restricted) Fisher-Wright bounds for $g \in \mathcal{G}^0$

The starting point is the fact (see Figure 8) that for each $\varepsilon \in \left(0, \frac{1}{3}\right]$ a given $g \in \mathcal{G}^0$ can be bounded as follows:

$$g^\varepsilon := b^\varepsilon f^\varepsilon \le g \le b^1 f \tag{66}$$

for some constants $b^\varepsilon, b^1 > 0$ where

$$f^\varepsilon(r) := (r-\varepsilon)^+(1-\varepsilon-r)^+, \qquad 0 \le r \le 1, \quad 0 \le \varepsilon \le \tfrac{1}{3}, \tag{67}$$

(recall that g is strictly positive on $(0,1)$ and Lipschitz). Here g^ε belongs to the more general set $\mathcal{G} \supset \mathcal{G}^0$ of diffusion coefficients introduced in Definition 5.3. We call g^ε a *restricted* Fisher-Wright diffusion coefficient. It is needed for the case of a diffusion coefficient g with a vanishing derivative at a boundary point of $[0,1]$ (as for instance in the Ohta-Kimura diffusions case $g = f^2$).

The *moment comparison principle* of [CFG96] we want to exploit says, roughly speaking, that larger diffusion coefficients lead to larger moments of X :

Proposition 5.7 (comparison of mixed moments) *For $\varepsilon \in \left[0, \frac{1}{3}\right]$, let positive constants b^ε and b^1 be given. Assume that $g \in \mathcal{G}$ satisfies (66) with f^ε defined in (67). Then the following higher moment inequalities hold:*

$$\mathbb{E}_z^{g^\varepsilon} X_{t_1}^{\psi^1} \cdots X_{t_k}^{\psi^k} \;\leq\; \mathbb{E}_z^{g} X_{t_1}^{\psi^1} \cdots X_{t_k}^{\psi^k} \;\leq\; \mathbb{E}_z^{b^1} X_{t_1}^{\psi^1} \cdots X_{t_k}^{\psi^k} \tag{68}$$

for all $z \in [0,1]^\Xi$, $k \geq 1$, $\psi^1, ..., \psi^k \in \Psi$, and $0 \leq t_1 \leq ... \leq t_k$.

Next we want to justify more formally why g^ε is called a *restricted* Fisher-Wright diffusion coefficient. For $0 \leq \varepsilon \leq \frac{1}{3}$ set

$$L_\varepsilon r := \tfrac{r-\varepsilon}{1-2\varepsilon}\,, \qquad 0 \leq r \leq 1, \tag{69}$$

which gives a map

$$L_\varepsilon : [0,1] \longmapsto \left[-\tfrac{\varepsilon}{1-2\varepsilon}\,, \tfrac{1-\varepsilon}{1-2\varepsilon}\right] =: I_\varepsilon \subseteq [-1,2]. \tag{70}$$

Applying coordinate-wise, L_ε can be considered as an affine mapping $L_\varepsilon : [0,1]^\Xi \mapsto I_\varepsilon^\Xi$.

Lemma 5.8 (restricted Fisher-Wright) *If z belongs to the set $[\varepsilon, 1-\varepsilon]^\Xi$ of restricted states, then under $\mathbb{P}_z^{g^\varepsilon}$, the transformed process $L_\varepsilon X$ has the law $\mathbb{P}_{L_\varepsilon z}^{b^\varepsilon}$ on $[0,1]^\Xi$.*

In fact,

$$g^\varepsilon(r) = (1-2\varepsilon)^2\, b^\varepsilon f\left(\tfrac{r-\varepsilon}{1-2\varepsilon}\right), \qquad r \in [\varepsilon, 1-\varepsilon].$$

Consequently, for truncated initial states, $L_\varepsilon X$ is an interacting Fisher-Wright diffusion on $[0,1]^\Xi$ with diffusion parameter b^ε.

5.4 Universality conclusion

The purpose of this subsection is to demonstrate how coupling and comparison are combined to prove universality statements on interacting diffusions, that is to reduce proofs to the Fisher-Wright case starting with a product initial law. The latter case amounts using the time-space duality relation (65) and the approximation Proposition 2.2 to showing a limit assertion on coalescing random walks η with immigration.

Theorem 6 (universality conclusion) *Fix natural numbers $k \geq 0$, $n_i \geq m_i \geq 0$, $0 \leq i \leq k$. For $t > 1$, let time points $s_0(t) < \cdots < s_{k+1}(t)$ be given such that $s_{i'} - s_i \to \infty$ as $t \to \infty$ if $i' > i$. Furthermore, for $0 \leq i \leq k$, pick*

$$\psi^i(t) \in \Psi, \quad \varphi^i(t) \in \Phi, \quad \psi^i(t) \wedge 1 = \varphi^i(t), \quad \|\psi^i(t)\| \equiv n_i\,, \quad \|\varphi^i(t)\| \equiv m_i\,.$$

Assume the coalescing random walk η with immigration satisfies

$$\mathbf{E}_{s_0(t),\ldots,s_k(t)}^{\varphi^0(t),\ldots,\varphi^k(t)}\, \theta^{\|\eta(s_{k+1}(t))\|} \underset{t\to\infty}{\longrightarrow} \text{ some } A^{m_0,\ldots,m_k}(\theta), \qquad 0 < \theta < 1, \tag{71}$$

(where $A^{m_0,\ldots,m_k}(\theta)$ is independent of $b > 0$). Then, for every g in $\mathcal{G}^0$ and $\mu \in \mathcal{T}_\theta$, $0 < \theta < 1$, the corresponding interacting diffusion X satisfies

$$\mathbb{E}_\mu^g X_{s_{k+1}(t)-s_0(t)}^{\psi^0(t)} \cdots X_{s_{k+1}(t)-s_k(t)}^{\psi^k(t)} \underset{t\to\infty}{\longrightarrow} A^{m_0,\ldots,m_k}(\theta). \tag{72}$$

Proof *Step* 1° We show that without loss of generality, in (72) we may restrict to the Fisher-Wright case $g = bf,\ b > 0$. Indeed, put $0 < \varepsilon \le \frac{1}{3}$ and, for the given $g \in \mathcal{G}^0$, choose $b^\varepsilon, b^1 > 0$ such that $g^\varepsilon = b^\varepsilon f^\varepsilon \le g \le b^1 f$. Apply the moment comparison (68) to see first that we have to deal only with the lower bound

$$\mathbb{E}_\mu^{g^\varepsilon} X^{\psi^0(t)}_{s_{k+1}(t)-s_0(t)} \cdots X^{\psi^k(t)}_{s_{k+1}(t)-s_k(t)}$$

once we know (72) in the Fisher-Wright case.

By the truncation Lemma 5.5, except some uniform ε–error $O(\varepsilon)$, we can replace μ by the ε–truncated law $\mu^\varepsilon \in \mathcal{T}_{\theta^\varepsilon}$ with $\theta^\varepsilon \to \theta$ as $\varepsilon \to 0$.

Next we use that for fixed $m \ge 0$,

$$x^\psi = \left(L_\varepsilon x\right)^\psi + O(\varepsilon) \quad \text{as } \varepsilon \downarrow 0, \tag{73}$$

uniformly in $x \in [0,1]^\Xi$ and $\psi \in \Psi$ with $\|\psi\| = m$ (the maps L_ε had been defined in (69)). Therefore from X we may switch to $L_\varepsilon X$, again except some uniform ε–error $O(\varepsilon)$. But by Lemma 5.8, the transformed process $L_\varepsilon X$ is an interacting Fisher-Wright diffusion on $[0,1]^\Xi$ with diffusion parameter b^ε. Hence,

$$\mathbb{E}_{\mu^\varepsilon}^{g^\varepsilon} L_\varepsilon X^{\psi^0}_{s_{k+1}-s_0} \cdots L_\varepsilon X^{\psi^k}_{s_{k+1}-s_k} = \mathbb{E}_{L_\varepsilon \mu^\varepsilon}^{b^\varepsilon} X^{\psi^0}_{s_{k+1}-s_0} \cdots X^{\psi^k}_{s_{k+1}-s_k}$$

with $L_\varepsilon \mu^\varepsilon$ the law of $L_\varepsilon z$ if z is distributed according to μ^ε. Since the limit in (72) (or (71)) is continuous in $\theta \in (0,1)$, and $(\theta^\varepsilon - \varepsilon)/(1-2\varepsilon) \longrightarrow \theta$ as $\varepsilon \to 0$, we get the *same* limit $A^{m_0,\dots,m_k}(\theta)$ for the lower bound after $\varepsilon \to 0$, once we know (72) in the Fisher-Wright case. This proves the claimed reduction to the Fisher-Wright case.

Step 2° Since we are now in the Fisher-Wright case $g = bf,\ b > 0$, we may apply the successful coupling Lemma 5.6, to reduce (72) to *product initial laws* $\mu \in \mathcal{T}_\theta$, that is if $\mu \in \mathcal{T}_\theta$ has the form

$$\mu = \textstyle\prod_{\xi \in \Xi} \mu_\xi, \qquad \displaystyle\int \mu_\xi(dr)\, r = \theta. \tag{74}$$

On the other hand, by the time-space duality (65), we rewrite the l.h.s. of (72):

$$\int \mu(\mathrm{d}z)\ \mathbb{E}_z^b X^{\psi^0}_{s_{k+1}-s_0} \cdots X^{\psi^k}_{s_{k+1}-s_k} = \int \mu(\mathrm{d}z)\ \mathbf{E}^{\psi^0,\dots,\psi^k}_{s_0,\dots,s_k}\, z^{\vartheta(s_{k+1})}.$$

By the approximation Proposition 2.2 we may replace this r.h.s. by

$$\int \mu(\mathrm{d}z)\ \mathbf{E}^{\varphi^0,\dots,\varphi^k}_{s_0,\dots,s_k}\, z^{\eta(s_{k+1})} = \mathbf{E}^{\varphi^0,\dots,\varphi^k}_{s_0,\dots,s_k}\, \theta^{\|\eta(s_{k+1})\|},$$

where we used the fact that by our reduction the initial state has i.i.d. components with expectation θ. However, this is the l.h.s. of (71), and the proof is finished. □

6 Limit statements for interacting diffusions

The purpose of this section is to use the results of the Sections 2–5 to prove the Theorems 1–3 of the introduction. To warm up we first want to prove the noise property of a single component process mentioned in § 1.3 above (p. 9).

6.1 Noise property of a single component process

Fix $\theta \in (0,1)$ and recall that $\mathcal{T}_\theta$ denotes the set of all *shift ergodic* initial distributions with density θ.

Proposition 6.1 (stationary 0-1–noise of components) *Let* $\mu \in \mathcal{T}_\theta$ *and* $g \in \mathcal{G}^0$. *Fix* $\xi \in \Xi$, $k \geq 0$ *and* $0 < \beta_1 < \ldots < \beta_{k+1}$. *Then*

$$\mathbb{P}^g_\mu \Big\{ \Big[X_\xi(\beta_1 t), \ldots, X_\xi(\beta_{k+1} t) \Big] \in \cdot \Big\} \underset{t\to\infty}{\Longrightarrow} \big[(1-\theta)\delta_0 + \theta\delta_1 \big]^{k+1}.$$

Consequently, here the limiting process is independent in each point and stationary. Actually this property essentially follows from the fact that in the Fisher-Wright dual, namely the delayed coalescing random walk with immigration *no particle will interact* in the present scaling regime.

Proof Fix μ, g, ξ, k and $\beta_1, \ldots, \beta_{k+1}$ as in the lemma. Without loss of generality, assume $\beta_{k+1} = 1$, $t = N^T$, and $\xi = 0$. We apply the *method of moments*. Choose integers $n_1, \ldots, n_{k+1} \geq 1$. It suffices to show that

$$\mathbb{E}^g_\mu X_0^{n_{k+1}}(\beta_{k+1} N^T) \cdots X_0^{n_1}(\beta_1 N^T) \underset{T\to\infty}{\longrightarrow} \theta^{k+1}. \tag{75}$$

According to Remark 5.2, one can interchange the antiton order in the time-space duality Proposition 5.1. Applied to Theorem 6 with $\psi^i = n_i \delta^0$ and $s_0(T) \equiv 0$, this means that (75) will follow if for the coalescing random walk η:

$$\mathbf{E}^{\delta^0, \ldots, \delta^0}_{\sigma_{k+1}(T), \ldots, \sigma_1(T)} \, \theta^{\|\eta(N^T)\|} \underset{T\to\infty}{\longrightarrow} \theta^{k+1} \tag{76}$$

where $\sigma_i(T) := N^T - \beta_i N^T = (1-\beta_i) N^T, \ 1 \leq i \leq k+1$.

However, by the walk speed Lemma 2.6 at p. 20, at the first immigration time $\sigma_k(T) = (1 - \beta_k) N^T$ the initial particle (starting at time $\sigma_{k+1}(T) = 0$ at 0) is located in $\Xi[T, 1]$ (defined in (35)) with probability approaching one as $T \to \infty$. Consequently, it is of order $T + o(T)$ away from the next immigrating particle. In the time $\sigma_{k-1}(T) - \sigma_k(T) = (\beta_k - \beta_{k-1}) N^T$ until the next immigration, the resulting difference walk moves away again of order $T + o(T)$. Hence these particles will not meet and will both be away from the origin. Continuing to argue in this way we see that $\|\eta(N^T)\| = k+1$ with probability converging to 1 as $T \to \infty$. Hence (76) holds, finishing the proof. □

6.2 Time-space thinned-out systems and Proof of Theorem 1

We shall deduce Theorem 1 from a more general statement allowing to simultaneously look at several component processes, which is interesting in its own despite the cumbersome notation one has to use.

In the following, for $\varphi \in \Phi$, we also write X_φ for the family $\{X_\xi \, ; \, \xi \in \Xi, \ \varphi_\xi > 0\}$. Recall the Definition 2.8 at p. 21 on spreading multi-colonies.

Theorem 7 (time-space thinned-out systems) *Assume* $\mu \in \mathcal{T}_\theta$, $\theta \in (0,1)$, *and* $g \in \mathcal{G}^0$. *Fix a constant* $c \geq 1$, *and assume the* $\alpha{\leq}\beta$*–Condition* 3.4 *at p.* 25 *with* $\beta_0 := 0$ *and* $\beta_{k+1} := 1$. *Consider spreading multi-colonies*

$$\varphi^i = \varphi^i(T) \in \overline{\Phi}_T \big[\underline{\alpha}_i, \underline{m}_i \, ; \, c \big], \qquad T > 1, \quad 0 \leq i \leq k.$$

(a) **(convergence of spreading families of components)** *Then the distributions*

$$\mathbb{P}_\mu^g\Big\{\Big[X_{\varphi^0(T)}\big(N^T - N^{\beta_0 T}\big), ..., X_{\varphi^k(T)}\big(N^T - N^{\beta_k T}\big)\Big] \in \cdot\Big\} \tag{77}$$

have a limit law as $T \to \infty$*, denoted by* $L = L_{\underline{\beta}}^{\underline{\underline{\alpha}};\underline{\underline{m}}}$*, concentrated on* $\{0,1\}^{|\underline{\underline{m}}|}$*. Here* $\underline{\underline{\alpha}}$*,* $\underline{\underline{m}}$ *and* $\underline{\beta}$ *are defined as in* (50) *at p. 26, and* $|\underline{\underline{m}}| := \sum_{i,j} m_{i,j}$*. The limit distribution* L *depends on the initial density* θ *but is otherwise independent of the "input data"* N, a, g, μ *of* X*.*

(b) **(characterization of the limit laws)** *This family of limit laws* $L_{\underline{\beta}}^{\underline{\underline{\alpha}};\underline{\underline{m}}}$ *is characterized by the fact that the probability for all components to be equal to* 1 *is given by*

$$\mathbf{E}^\theta\Big[\prod_{i=0}^{k}\prod_{j=0}^{\ell_i}\Big(\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha_{i,j})\Big)^{m_{i,j}}\Big] \tag{78}$$

with $\widetilde{\mathbf{Y}}^\theta$ *the transformed Fisher-Wright tree of Definition* 1.9 *at p.* 7*.*

In particular, the distribution of the limit array is a mixture of Bernoulli product laws: First realize the "weights"

$$\Big\{\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha_{i,j});\ \ 0 \le i \le k,\ \ 0 \le j \le \ell_i\Big\}$$

according to the distribution $\mathbf{P}^\theta$ *of the transformed Fisher-Wright tree* $\widetilde{\mathbf{Y}}^\theta$*, and then form the product laws with marginals*

$$\big(1 - \widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha_{i,j})\big)\delta_0 + \widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha_{i,j})\,\delta_1\,, \qquad 0 \le i \le k, \quad 0 \le j \le \ell_i\,.$$

Remark 6.2 From the point of view of the interacting diffusion X, the $\alpha{\le}\beta$–Condition 3.4 at p. 25 just reflects the *natural range of growth of clusters.* ◇

Proof (a) Set $s_i(T) := N^{\beta_i T}$, $0 \le i \le k$. In order to apply the *method of moments,* take "T–independent multiples" $\psi^i(T)$ of $\varphi^i(T)$, that is $\psi^i(T) \in \Psi$ satisfying $\psi^i(T) \wedge 1 = \varphi^i(T)$, and where the multiplicities $\psi^i_\zeta(T) > 0$ are independent of T. We want to show that

$$\mathbb{E}_\mu^g \prod_{i=0}^{k} X^{\psi^i(T)}\big(N^T - s_i(T)\big) \xrightarrow[T\to\infty]{} \sum_{n=1}^{\infty} \mathcal{P}\Big(\widetilde{\Lambda}_{\underline{\beta}}^{\underline{\underline{\alpha}};\underline{\underline{m}}}(1) = n\Big)\theta^n \tag{79}$$

with $\widetilde{\Lambda}$ the ensemble of log-coalescents with immigration of Definition 3.5. Since the r.h.s. is independent of b, according to the universality conclusion Theorem 6 it suffices to show (79) with the l.h.s. replaced by

$$\mathbf{E}^{\varphi^0(T),...,\varphi^k(T)}_{s_0(T),...,s_k(T)}\,\theta^{\|\eta(N^T)\|}$$

(recall that $\psi^i \wedge 1 = \varphi^i$). But then by the *scaling limit Theorem* 4 at p. 27 on coalescing random walks with immigrating multi-colonies the claim (79) follows. Hence, the limit law $L_{\underline{\beta}}^{\underline{\underline{\alpha}};\underline{\underline{m}}}$ exists and is concentrated on $\{0,1\}^{|\underline{\underline{m}}|}$ since the limit expression in (79) is independent of the orders $\psi^i_\zeta(T) > 0$ of moments at the l.h.s. of (79).

(b) Using additionally the characterization (59) at p. 31 of the duality Theorem 5, we get the limiting probability of all components to be 1 as claimed in (78). This finishes the proof. □

Proof of Theorem 1 Specialize the assumptions in Theorem 7 as follows: $\ell_i \equiv 0$, $m_{i,0} \equiv m_i \leq 1$, $\alpha_{i,0} \equiv 0$ and $\varphi^i \equiv \delta^\xi$ if $m_i = 1$. Then the claims (a) and (b) of Theorem 7 imply the corresponding ones in Theorem 1, as in particular already explained in Example 4.1 at p. 31.

The marginal laws $\mathcal{L}(U_\beta^\infty)$ are given by the basic ergodic theorem (14). Since we can map $r \mapsto 1-r$, we may fix our attention on the case $U_0^\infty = \partial = 1$. We need to show that

$$\mathbb{P}\Big(U_0^\infty = 1,\ h_U \geq \beta\Big) = P\Big(Y_\tau^\theta = 1,\ \tau \leq \log(1/\beta)\Big), \qquad 0 < \beta < 1. \tag{80}$$

The l.h.s. coincides with the monotone limit

$$\lim_{n\to\infty} \mathbb{P}\Big(U_{k2^{-n}\beta}^\infty = 1,\ \ 0 \leq k \leq 2^n\Big),$$

which by the identity in (b) is equal to

$$\lim_{m\to\infty} E\big(Y_t^\theta\big)^m \quad \text{with} \quad t := \log(1/\beta).$$

If we restrict the latter expectation additionally to the event from the r.h.s. of (80), then $Y_t^\theta = 1$, and we actually arrive at the r.h.s. of (80). The remaining part of the expectation can be bounded from above by

$$E\Big\{\big(Y_t^\theta\big)^m;\ Y_t^\theta < 1\Big\},$$

which converges to 0 as $m \to \infty$, by bounded convergence. This verifies (80), and shows that $\big[U_0^\infty, h_U\big]$ has the claimed law. But combined with (b), the remaining claim of (c) follows immediately. This finishes the proof. □

6.3 Time-space thinned-out systems: Proof of Theorem 3

1° (*convergence and characterization*) Fix $\mu \in \mathcal{T}_\theta$, $0 < \theta < 1$, natural numbers $k, m_0, ..., m_k \geq 0$, and constants $1 > \beta_1 > \cdots > \beta_k \geq \alpha > 0 = \beta_0$. For $i \in \{0, ..., k\}$, consider $\chi^i \in \Phi$ with $\|\chi^i\| = m_i$. For $T > 1$, write $s_i(T) := \sum_{1 \leq i' \leq i} N^{\beta_{i'} T}$ and

$$\varphi^i = \varphi^i(T) := \mathcal{S}_{[\alpha T]}^{-1}\chi^i := \textstyle\sum_{\xi:\ \chi_\xi^i > 0} \delta^{\mathcal{S}_{[\alpha T]}^{-1}\xi}$$

for the spread-out population, giving the particles of χ an asymptotic distance αT, as in the thinning procedure of Definition 1.14 (a). As in the proof of Theorem 7, apply the method of moments, and take "T–independent multiples" $\psi^i \in \Psi$ of φ^i. In order to determine the limit in law as $T \to \infty$ of the array of variables

$$\Big\{W_{\xi,i}^{\underline{\beta},\alpha,T};\ \ \xi \in \operatorname{supp} \chi_\xi^i\,,\ \ 0 \leq i \leq k\Big\} \tag{81}$$

we look at the moments

$$\mathbb{E}_\mu^g \textstyle\prod_{i=0}^k X^{\psi^i(T)}\big(N^T - s_i(T)\big). \tag{82}$$

According to the universality conclusion Theorem 6, we need to study

$$\mathbf{E}_{s_0(T),\ldots,s_k(T)}^{\varphi^0(T),\ldots,\varphi^k(T)}\, \theta^{\|\eta(N^T)\|}.$$

By Proposition 3.10 at p. 29, this generating function in θ converges as $T \to \infty$ to the corresponding one of the ensemble $\widetilde{\Lambda}^{[\alpha,\dots,\alpha];[m_k,\dots,m_0]}_{[\beta_k,\dots,\beta_1]}$ of log-coalescents without immigration. But according to the duality Theorem 5, the latter generating function is given by

$$\mathbf{E}^\theta \prod_{i=0}^k \bigl(\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\alpha)\bigr)^{m_i}. \tag{83}$$

Hence, the moments (82) have the limit (83), and we conclude that the limiting field $W^{\underline{\beta},\alpha,\infty}$ exists. Moreover, since (83) is independent of the orders $\psi^i_\xi(T) > 0$ of the moments, the limit $W^{\underline{\beta},\alpha,\infty}$ is $\{0,1\}^{\Xi\times\{0,\dots,k\}}$–valued. This implies claims (a) and (b) of the theorem.

2° (*a moment estimate*) Consider (83), and condition on $\widetilde{\mathcal{F}}(\beta_k)$. Then all factors become independent, and we can switch to a product of conditional moments. Moreover, by Jensen's inequality, these conditional moments can be bounded below by corresponding powers of first moments. But by the martingale property of the Fisher-Wright diffusion, these expectations can be computed arriving altogether at the lower estimate

$$\mathbf{E}^\theta \prod_{i=0}^k \bigl(\widetilde{\mathbf{Y}}^\theta{}_{\beta_i}(\beta_k)\bigr)^{m_i}$$

for (83). Since $\widetilde{\mathbf{Y}}^\theta{}_{\beta_k}(\beta_k) = \widetilde{\mathbf{Y}}^\theta{}_0(\beta_k)$, we actually reduced (83) by one factor. Hence, by induction, we will end up with the lower estimate $\theta^{m_0+\cdots+m_k}$ for (83). Consequently, from (24),

$$\mathbb{E} \prod_{\xi\in\operatorname{supp}\chi^i_\xi,\ 0\le i\le k} W^{\underline{\beta},\alpha,\infty}_{\xi,i} \;\ge\; \theta^{m_0+\cdots+m_k} \quad \text{where} \quad \theta \equiv \mathbb{E}W^{\underline{\beta},\alpha,\infty}_{\xi,i}\,. \tag{84}$$

3° (*association*) By definition, a countable family of variables is *associated* if every two non-decreasing functions of this family, depending only on finitely many components and being square integrable with respect to the law of the family, are non-negatively correlated. Our last formula implies that for events C, D of the form

$$\bigl\{W^{\underline{\beta},\alpha,\infty}_{\xi,i} = 1 \ \text{ for } \ \xi\in A \text{ and } i\in B\bigr\}, \tag{85}$$

where A, B are finite subsets of Ξ and $\{0,\dots,k\}$, respectively,

$$\mathbb{P}(C\cap D) \;\ge\; \mathbb{P}(C)\,\mathbb{P}(D).$$

According to [Lin88] it suffices to have this property for all increasing events (depending only on finitely many components). Since the variables are 0-1–valued, all increasing events are of the form (85), and the needed property follows. Hence, the limit field $W^{\underline{\beta},\alpha,\infty}$ is associated.

4° (*representation*) The claimed representation immediately follows from the characterization (24) combined with the trapping property (13), finishing the proof. □

6.4 Spatial ball averages: Proof of Theorem 2

Fix $g\in\mathcal{G}^0$, $\mu\in\mathcal{T}_\theta$, $\theta\in(0,1)$, and $0\le\alpha<1$. Recall the definition (19) of $V^{\alpha,T}$. If $\alpha=0$, then $V^{\alpha,T}=U^T$. Hence, for the proof of the convergence statement we may restrict to $0<\alpha<1$. By spatial homogeneity, we may also set $\xi=0$.

1° (*asymptotic moment formula*) The process $V^{\alpha,T}$ takes on values in $[0,1]$ only, thus we can again use the *method of moments*. Fix $k, m_0,\dots,m_k\ge 0$ and $0=:\beta_0<\cdots<\beta_{k+1}:=1$, and consider

$$\mathbb{E}^g_\mu \bigl(V^{\alpha,T}_{\beta_0}\bigr)^{m_0}\cdots\bigl(V^{\alpha,T}_{\beta_k}\bigr)^{m_k}. \tag{86}$$

In order to evaluate this moment, we use definition (19) of $V^{\alpha,T}$ as average over components X_ξ of X to get

$$= N^{-M_k[\alpha T]} \sum_{\xi(1),\dots,\xi(M_k)\in\Xi_{\alpha T}} \mathbb{E}^g_\mu\Big[\prod_{i=0}^{k}\prod_{j_i=M_{i-1}+1}^{M_i} X_{\xi(j_i)}(N^T - N^{\beta_i T})\Big] \tag{87}$$

where $\Xi_r := \{\xi \in \Xi;\ \|\xi\| \le r\}$ and $M_i := m_0 + \cdots + m_i$, $i = -1, 0, ..., k$. We may assume that $M_k \ge 1$. Set $s_i(T) := N^{\beta_i T}$, $0 \le i \le k$.

2° (*restriction of the range of summation*) Asymptotically as $T \to \infty$ we may restrict the range of summation in (87) requiring additionally

$$\|\xi(j)\| \wedge \|\xi(j) - \xi(j')\| \ge [\alpha T] - \ell(\alpha T), \qquad j \ne j', \quad 1 \le j, j' \le M_k \tag{88}$$

(recall the definition (34) of $\ell(r)$ at p. 20). Roughly speaking, we sum only over labels with absolute and relative speed α. To justify this restriction, first note that all terms of the sum in (87) are uniformly bounded by 1. Moreover, the number of labels excluded this way is bounded by

$$C(\underline{m})\, N^{M_k[\alpha T] - \ell(\alpha T)} = o\Big(N^{M_k[\alpha T]}\Big) \quad \text{as } T \to \infty,$$

where the combinatorial constant $C(\underline{m})$ only depends on $\underline{m} := [m_0, ..., m_k]$. In fact, we have $C(\underline{m})\, N^{(M_k-1)[\alpha T]}$ possibilities to fix $\xi(j')$, then the coordinates $\xi_i(j)$ of $\xi(j)$ with $i \ge [\alpha T] - \ell(\alpha T)$ have to be 0 or coincide with the corresponding ones of $\xi(j)$ in order to *violate* the inequality in (88); this gives further $[\alpha T] - \ell(\alpha T)$ possibilities.

3° (*convergence*) Set $\varphi^i := \delta^{\xi(M_{i-1}+1)} + \cdots + \delta^{\xi(M_i)}$, $0 \le i \le k$, with the $\xi(j)$ of the range of summation in (87) but with the restriction (88). Note that $\varphi^0 + \cdots + \varphi^k$ belongs to $\Phi_T[\alpha, M_k, 1]$ for all sufficiently large T. (We applied the Definition 2.8 of spreading multi-colonies, specialized to a single colony.) Then a typical term in the sum in (87) can be written as

$$\mathbb{E}^g_\mu \textstyle\prod_{i=0}^k X^{\varphi^i(T)}\big(s_{k+1} - s_k\big). \tag{89}$$

In order to calculate the limit of (89) as $T \to \infty$ which then gives the limit of (87), we want to apply the universality conclusion Theorem 6. Therefore we look at

$$\mathbf{E}^{\varphi^0,\dots,\varphi^k}_{s_0,\dots,s_k}\, \theta^{\|\eta(N^T)\|}. \tag{90}$$

Then by Proposition 3.9 at p. 28, as $T \to \infty$ we get the following limit for the generating function (90):

$$\textstyle\sum_{n=0}^\infty \mathcal{P}\Big(\widetilde{\Lambda}^{\alpha,\dots,\alpha,\,M_{J-1},\,m_J,\dots,m_k}_{\beta_J,\dots,\beta_k}(1) = n\Big)\theta^n, \tag{91}$$

with J defined in (21). Consequently, (89) hence (86) converges to (91) by the universality conclusion Theorem 6. This shows that indeed a process $V^{\alpha,\infty}$ on $[0,1)$ exists such that (20) holds, and the statement (a) of Theorem 2 is proved.

4° (*limiting moments*) By the duality relation (59) of Theorem 5, the limiting moments of (86) as $T \to \infty$, coincide with (91) and hence fulfill the identity claimed in (b).

5° (*marginals*) Specializing the moment formula of (b) to $k = 0$ (implying $J = 1$), we immediately see that the random variables $V^{\alpha,\infty}_\beta$ and $\widetilde{\mathbf{Y}^\theta}_0(\alpha) = \widetilde{Y^\theta}(\alpha)$ coincide in law. Thus, by (10) we get the marginals as claimed in the beginning of (c).

6° (*holding time and conditional noise*) From the case $J = k+1$ in the moment identity of (b) we conclude that the limit process $V^{\alpha,\infty}$ is constant at least on the interval $[0,\alpha)$ (where the constant is random with law $\widetilde{Q^\theta}_\alpha$). That is, $h_V \geq \alpha$. Moreover, putting $J=1$ in the moment formula of (b), we see that for $0<\alpha\leq\beta_1<\cdots\beta_k<1$,

$$\left[V_0^{\alpha,\infty}, V_{\beta_1}^{\alpha,\infty}, ..., V_{\beta_k}^{\alpha,\infty}\right] \stackrel{\mathcal{L}}{=} \left[\widetilde{\mathbf{Y}^\theta}_0(\alpha), \widetilde{\mathbf{Y}^\theta}_{\beta_1}(\alpha), ..., \widetilde{\mathbf{Y}^\theta}_{\beta_k}(\alpha)\right]. \tag{92}$$

By definition of the (transformed) Fisher-Wright tree $\widetilde{\mathbf{Y}^\theta}$, given the trunk $\widetilde{\mathbf{Y}^\theta}_0$, the (backward) branches $\widetilde{\mathbf{Y}^\theta}_{\beta_1}, ..., \widetilde{\mathbf{Y}^\theta}_{\beta_k}$ are independent. Hence, if we condition the r.h.s. of (92) to the trunk $\widetilde{\mathbf{Y}^\theta}_0$ and restrict additionally to $\beta_k \leq \widetilde{\tau}$, then

$$\widetilde{\mathbf{Y}^\theta}_0(\alpha) = \widetilde{\mathbf{Y}^\theta}_{\beta_1}(\alpha) = \cdots = \widetilde{\mathbf{Y}^\theta}_{\beta_k}(\alpha),$$

by (13). Hence, $\widetilde{\mathbf{Y}^\theta}_\beta(\alpha) = \widetilde{\mathbf{Y}^\theta}_0(\alpha)$ whenever $\alpha\leq\beta\leq\widetilde{\tau}$. (Actually, for such a statement one has to extend the definition of the trees, switching to an uncountable collection of branches.) That is, the holding time of the "process" $\beta\mapsto\widetilde{\mathbf{Y}^\theta}_\beta(\alpha)$ on $[\alpha,1)$ is at least $\widetilde{\tau}\vee\alpha$. Our aim is to demonstrate that this holding time is actually exactly $\widetilde{\tau}\vee\alpha$.

Given the trunk and restricting to $(\alpha\vee\widetilde{\tau})<\beta_1$, the termination positions

$$\left[\widetilde{\mathbf{Y}^\theta}_{\beta_1}(\beta_1), ..., \widetilde{\mathbf{Y}^\theta}_{\beta_k}(\beta_k)\right] = \left[\widetilde{\mathbf{Y}^\theta}_0(\beta_1), ..., \widetilde{\mathbf{Y}^\theta}_0(\beta_k)\right] = \left[\widetilde{Y^\theta}(\beta_1), ..., \widetilde{Y^\theta}(\beta_k)\right] =: \left[\theta_1, ..., \theta_k\right]$$

of the (conditional) independent branches $\widetilde{\mathbf{Y}^\theta}_{\beta_1}, ..., \widetilde{\mathbf{Y}^\theta}_{\beta_k}$ are interior points of $[0,1]$ implying that the corresponding branches are non-degenerate. More precisely, given θ_i, by homogeneity as in the property (b) of Lemma 4.3 at p. 32, and then switching to the trunk we get

$$\widetilde{\mathbf{Y}^\theta}_{\beta_i}(\alpha) \stackrel{\mathcal{L}}{=} \widetilde{\mathbf{Y}^{\theta_i}}_1(\alpha/\beta_i) \stackrel{\mathcal{L}}{=} \widetilde{\mathbf{Y}^{\theta_i}}_0(\alpha/\beta_i) = \widetilde{Y^{\theta_i}}(\alpha/\beta_i), \tag{93}$$

which by definition has the law $\widetilde{Q^{\theta_i}}_{\alpha/\beta_i}$. This verifies that the conditional distribution

$$\mathcal{L}\left\{\left[\widetilde{\mathbf{Y}^\theta}_{\beta_1}(\alpha), ..., \widetilde{\mathbf{Y}^\theta}_{\beta_k}(\alpha)\right] \,\middle|\, (\alpha\vee\widetilde{\tau})<\beta_1\right\}$$

of the "section" of the transformed Fisher-Wright tree is just a mixture of product laws as claimed in (c).

It remains to show the statement (c2). Given the trunk, we choose an $\varepsilon>0$ such that $(\alpha\vee\widetilde{\tau})+\varepsilon<1$. Then, abbreviating $\widetilde{\mathbf{Y}^\theta}_0(\alpha) =: r$,

$$\mathbf{P}^\theta\left\{r = \widetilde{\mathbf{Y}^\theta}_{\beta_1}(\alpha) = \cdots = \widetilde{\mathbf{Y}^\theta}_{\beta_k}(\alpha) \,\middle|\, \widetilde{\mathbf{Y}^\theta}_0;\ (\alpha\vee\widetilde{\tau})<\beta_1<\cdots\beta_k<(\alpha\vee\widetilde{\tau})+\varepsilon\right\}$$
$$= \textstyle\prod_{i=1}^k \widetilde{Q^{\theta_i}}_{\alpha/\beta_i}(\{r\}),$$

which equals 0 if $\widetilde{\tau}<\alpha$ is satisfied, and converges to 0 as $k\to\infty$ otherwise. Thus, the holding time of $\beta\mapsto\widetilde{\mathbf{Y}^\theta}_\beta(\alpha)$ on $[\alpha,1)$ is exactly $\widetilde{\tau}\vee\alpha$, and by (92) we conclude that $h_V=\widetilde{\tau}\vee\alpha$ in law. This completes the proof of (c) and finishes the proof of Theorem 2 altogether. □

Acknowledgment We thank an anonymous referee for suggestions which led to an improvement of the exposition.

References

[CFG96] J.T. Cox, K. Fleischmann, and A. Greven. Comparison of interacting diffusions and application to their ergodic theory. *Probab. Theory Relat. Fields*, to appear 1996.

[CG83] J.T. Cox and D. Griffeath. Occupation time limit theorem for the voter model. *Ann. Probab.*, 11:876–893, 1983.

[CG85] J.T. Cox and D. Griffeath. Occupation times for critical branching Brownian motions. *Ann. Probab.*, 13(4):1108–1132, 1985.

[CG86] J.T. Cox and D. Griffeath. Diffusive clustering in the two dimensional voter model. *Ann. Probab.*, 14:347–370, 1986.

[CG94] J.T. Cox and A. Greven. Ergodic theorems for infinite systems of locally interacting diffusions. *Ann. Probab.*, 22(2):833–853, 1994.

[Cox89] J.T. Cox. Coalescing random walks and voter model consensus times on the torus in Z^d. *Ann. Probab.*, 17(4):1333–1366, 1989.

[Daw93] D.A. Dawson. Measure-valued Markov processes. In P.L. Hennequin, editor, *École d'été de probabilités de Saint Flour XXI–1991*, volume 1541 of *Lecture Notes in Mathematics*, pages 1–260, Berlin, Heidelberg, New York, 1993. Springer-Verlag.

[DF88] D.A. Dawson and K. Fleischmann. Strong clumping of critical space-time branching models in subcritical dimensions. *Stoch. Proc. Appl.*, 30:193–208, 1988.

[DG93] D.A. Dawson and A. Greven. Hierarchical models of interacting diffusions: Multiple time scale phenomena, phase transitions and pattern of clusterformation. *Probab. Theory Relat. Fields*, 96:435–473, 1993.

[DK96] P. Donelly and T.G. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.*, to appear 1996.

[EF96] S.N. Evans and K. Fleischmann. Cluster formation in a stepping stone model with continuous, hierarchically structered sites. WIAS Berlin, Preprint Nr. 187, 1995, *Ann. Probab.*, to appear 1996.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, New York, 1986.

[FG94] K. Fleischmann and A. Greven. Diffusive clustering in an infinite system of hierarchically interacting diffusions. *Probab. Theory Relat. Fields*, 98:517–566, 1994.

[Isc86] I. Iscoe. A weighted occupation time for a class of measure-valued critical branching Brownian motions. *Probab. Theory Relat. Fields*, 71:85–116, 1986.

[Kin82] J.F.C. Kingman. The coalescent. *Stoch. Proc. Appl.*, 13:235–248, 1982.

[Kle95] A. Klenke. Different clustering regimes in systems of hierarchically interacting diffusions. *Ann. Probab.*, to appear 1995.

[Lig85] T.M. Liggett. *Interacting Particle Systems.* Springer-Verlag, New York, 1985.

[Lin88] H. Lindqvist. Association of probability measures on partially ordered spaces. *J. Multivariate Analysis*, 26:111–132, 1988.

[SF83] S. Sawyer and J. Felsenstein. Isolation by distance in a hierarchically clustered population. *J. Appl. Probab.*, 20:1–10, 1983.

[Shi80] T. Shiga. An interacting system in population genetics. *J. Mat. Kyoto Univ.*, 20:213–242, 1980.

[SS80] T. Shiga and A. Shimizu. Infinite-dimensional stochastic differential equations and their applications. *J. Mat. Kyoto Univ.*, 20:395–416, 1980.

[Tav84] S. Tavaré. Line-of-descent and genealogical processes, and their applications in population genetics models. *Theoret. Population Biol.*, 26:119–164, 1984.

Weierstraß-Institut für Angewandte
Analysis und Stochastik (WIAS)
Mohrenstr. 39
D - 10117 Berlin, Germany
e-mail: fleischmann@wias-berlin.de

Mathematisches Institut
Universität Erlangen-Nürnberg
Bismarckstr. $1\frac{1}{2}$
D - 91054 Erlangen, Germany
e-mail: greven@mi.uni-erlangen.de