The Annals of Probability
1996, Vol. 24, No. 2, 660–697

DIFFERENT CLUSTERING REGIMES IN SYSTEMS OF HIERARCHICALLY INTERACTING DIFFUSIONS

BY ACHIM KLENKE

Universität Erlangen-Nürnberg

We study a system of interacting diffusions

$$dx_\xi(t) = \sum_{\zeta\in\Xi} a(\xi,\zeta)(x_\zeta(t) - x_\xi(t))\,dt + \sqrt{g(x_\xi(t))}\,dW_\xi(t) \qquad (\xi \in \Xi),$$

indexed by the hierarchical group Ξ, as a genealogical two genotype model [where $x_\xi(t)$ denotes the frequency of, say, type A] with hierarchically determined degrees of relationship between colonies. In the case of short interaction range it is known that the system clusters. That is, locally one genotype dies out. We focus on the description of the different regimes of cluster growth which is shown to depend on the interaction kernel $a(\cdot,\cdot)$ via its recurrent potential kernel. One of these regimes will be further investigated by mean-field methods. For general interaction range we shall also relate the behavior of large finite systems, indexed by finite subsets Ξ_n of Ξ, to that of the infinite system. On the way we will establish relations between hitting times of random walks and their potentials.

1. Introduction and main results.

Survey. In this paper we analyze the pattern of cluster formation in systems of interacting diffusions and study the behavior of large finite versus infinite systems of interacting diffusions. Our main point is to cover the full range of clustering models in a systematic way. So far the treatment of clustering phenomena has focused on particular interaction kernels [see Arratia (1982), Cox and Griffeath (1986) and Fleischmann and Greven (1994)] or the system has been studied after taking a parameter of the dynamics to a limit [see Dawson and Greven (1993a, b)]. In fact, we shall investigate the question whether the mean-field analysis of Dawson, Greven and Vaillancourt (1995) indeed yields the same result as when we take the objects describing the cluster formation for a given interacting system and then let the interaction parameter approach its limit.

At the same time we are likewise able to treat, in a systematic way, the question of how the behaviors of finite and infinite systems are related for systems on the hierarchical group for the whole class of models considered. For a treatment of the lattice case, see Cox and Greven (1990) and Cox, Greven and Shiga (1994).

Received August 1994; revised June 1995.
AMS 1991 *subject classification*. 60K35.
Key words and phrases. Interacting diffusions, infinite particle systems, coalescing random walks.

Introduction. We consider a system $\mathbb{X}(t) = (x_\xi(t))_{\xi\in\Xi}$ of linearly interacting diffusions on $[0,1]^\Xi$ defined as the solution of the system of stochastic differential equations (SSDE)

$$\begin{aligned} dx_\xi(t) = \sum_{\zeta\in\Xi} a(\xi,\zeta)(x_\zeta(t) - x_\xi(t))\,dt \\ + \sqrt{g(x_\xi(t))}\,dW_\xi(t) \qquad (\xi \in \Xi), \end{aligned} \tag{1.1}$$

indexed by the countable hierarchical group Ξ, where (W_ξ) are independent Brownian motions, $a(\cdot,\cdot)$ is the kernel of a random walk on Ξ and the *diffusion coefficient* g is assumed to satisfy

$$\begin{aligned} g\colon\ [0,1] \to [0,\infty[\text{ is Lipschitz continuous,} \\ g(x) = 0 \quad \text{iff } x \in \{0,1\}. \end{aligned} \tag{1.2}$$

Existence and uniqueness of the strong solution of (1.1) is assured by Shiga and Shimizu [(1980), Theorem 3.2].

The hierarchical group Ξ is defined by

$$\begin{aligned} \Xi := \bigl\{\xi = (\xi_m)_{m\in\mathbb{N}}\colon\ \xi_m \in \{0,\dots,N-1\}, \\ \xi_m \neq 0 \text{ only for finitely many } m\bigr\} \end{aligned} \tag{1.3}$$

with addition componentwise modulo N ($N = 2,3,\dots$ is some fixed parameter) and distance $\|\xi\| := \max\{k\colon\ \xi_k \neq 0\} \vee 0$. Of course Ξ carries the discrete topology, induced by the metrics $\|\cdot\|$. For $n \in \mathbb{N}_0 = \mathbb{N}\cup\{0\}$ we denote by Ξ_n the finite subgroup

$$\Xi_n := \{\xi \in \Xi\colon\ \|\xi\| \le n\}. \tag{1.4}$$

We restrict ourselves to the case where $a(\xi,\zeta)$ depends only on $\|\xi-\zeta\|$ and put for $k = \|\xi-\zeta\|$,

$$\begin{aligned} r_k &:= a(\xi,\zeta)R_k \quad \text{with} \\ R_k &:= \#\{\xi\in\Xi\colon\ \|\xi\| = k\} = (N - 1 + \mathbb{1}_0(k))N^{k-1}. \end{aligned} \tag{1.5}$$

This model was suggested by Sawyer (1976) to describe the evolution of gene frequencies. Think of Ξ as the site space, each site ξ containing a (large) colony of individuals. Then $x_\xi(t)$ represents the frequency of some fixed allele, say A, at site ξ and time t. By resampling, the frequency fluctuates at random, modelled by g. Additionally, the frequency may change by migration.

Here the spatial structure of the site space becomes important. The idea is that the colonies are organized according to different degrees of relationship: N colonies form a family, N families form a clan, N clans form a tribe and so on. Thus $\xi = (\xi_1,\xi_2,\xi_3,\dots)$ is the ξ_1th member of the ξ_2th family of the ξ_3th clan and so forth. We measure the degree of relationship between two colonies ξ and ζ by $\|\xi-\zeta\|$. If, for example $\|\xi-\zeta\| = 2$, then ξ and ζ are in the same clan, tribe and so forth, but in different families. The flow of migration between two colonies shall depend only on their degree of relationship. The total flow of migration from ξ to all relatives of degree k is r_k. It divides uniformly on all relatives of the same degree.

Here and in the following $\mu = \mathscr{L}^\mu(\mathbb{X}(0))$ is assumed to be in $\mathscr{M}_\theta$ (for some $\theta \in [0, 1]$) given by

$$\mathscr{M}_\theta = \{\mu\colon \mu \text{ is a spatially ergodic probabability measure on } \Xi \text{ with intensity } \theta = \langle \mu, x_0\rangle\}. \tag{1.6}$$

Note that spatial homogeneity of the starting measure is preserved under the dynamics.

It is known that $\mathbb{X}(t)$ clusters if $a(\cdot,\cdot)$ is recurrent. That is,

$$\mathscr{L}^\mu(\mathbb{X}(t)) \Rightarrow \theta\delta_{\mathbf{1}} + (1-\theta)\delta_{\mathbf{0}} \quad \text{as } t \to \infty, \tag{1.7}$$

where $\delta_{\mathbf{0}},\ \delta_{\mathbf{1}}$ denote the (unit) point masses on $\mathbf{0},\ \mathbf{1} \in [0,1]^\Xi$.

In the case where a transient, as opposed to (1.7), there is a family $(\nu_\theta | \theta \in [0,1])$ of invariant (under the dynamics) ergodic measures with intensity $\theta = \langle \nu_\theta, x_0\rangle$ such that for $\mu \in \mathscr{M}_\theta$,

$$\mathscr{L}^\mu(\mathbb{X}(t)) \Rightarrow \nu_\theta \quad \text{as } t \to \infty. \tag{1.8}$$

[See Cox and Greven (1994a), Theorems 1 and 2.]

Of special interest are the geometrical kernels $a_c,\ c > 1/N$, with $r_k = \vartheta_c (Nc)^{-k}$ [$\vartheta_c = (Nc-1)/Nc$ is the normalizing constant]. One can easily verify that a_c is transient iff $c < 1$ [see (2.31)].

NOTATION. We denote by $\mathscr{L}$ the law of a random variable, by $\Rightarrow$ weak convergence and let $\langle \mu, f\rangle = \int f\, d\mu$. Thus $\theta = \int x_0 \nu_\theta(dx)$.

Clustering in infinite systems. We are now led to the question of how fast the clusters grow in the case a recurrent. It has already been shown in the theory of interacting particle systems that this depends on the strength of interaction [see Bramson and Griffeath (1980) and Cox and Griffeath (1986)]. In our situation it depends on whether $c = 1$ or $c > 1$. In the first case, the so-called *diffusive* case, clusters grow at random speed. This was studied in great detail by Fleischmann and Greven (1994). However, we shall see that the diffusive case is not as singular as it seems at first glance by being sandwiched between $c < 1$ and $c > 1$. Namely, and this is our main point, it will be broadened to transition kernels such that $k \mapsto \log(N^k r_k)$ is slowly varying in a sense that will be made precise. Here the random speed of growth splits up into three regimes. We shall investigate this more closely in our Theorem 1. In contrast, in the case $c > 1$, clusters grow with a fixed deterministic speed and we shall study fluctuations in our Theorem 2.

In order to fix the notion of growing clusters we work with two concepts described in (i) and (ii) below.

(i) *Scaled systems.* In order to get a more detailed description of the clustering of (1.7) we want to compare sites with a distance growing in time. For a systematic treatment, however, we will also rescale the time by a monotone sequence (s_n), $s_n \uparrow \infty$, called the *time scale*. Thus for $n \in \mathbb{N}_0$ we consider sites of distance $f(n)$ at time s_n. The monotone function $f\colon \mathbb{N}_0 \to \mathbb{N}_0$, $f(n) \uparrow \infty$, is called *space scale*. To keep time continuous we introduce the "inverse" of (s_n):

$$n(t) = \sup\{n \in \mathbb{N}\colon\ s_n \le t\} \vee 0. \tag{1.9}$$

For f and (s_n) fixed the rescaled system ${}^f\mathbb{X}(t)$ is defined as follows.

Let the shift operators $S_k\colon \Xi \to \Xi$, $k = 0, 1, 2, \dots$, be defined by

$$S_k((\xi_m)_{m\in\mathbb{N}}) = (\xi_{m+k})_{m\in\mathbb{N}} \tag{1.10}$$

and let S_k^{-1} be a fixed right inverse. Now ${}^f\mathbb{X}(t) = ({}^f x_\xi(t))_{\xi\in\Xi}$ is defined by

$${}^f x_\xi(t) = x_\zeta(t), \quad \text{where } \zeta = S^{-1}_{f(n(t))}\xi. \tag{1.11}$$

(ii) *Block averages.* For $n \in \mathbb{N}$ let the nth block average be defined by

$$\begin{aligned} \Theta_n\colon\ [0,1]^\Xi &\to [0,1],\\ (x_\xi) &\mapsto N^{-n} \sum_{\xi\in\Xi_n} x_\xi. \end{aligned} \tag{1.12}$$

The block averages are to be thought of as macroscopic variables determining the behavior of the system up to a certain degree. So as to fully explore this concept we have introduced the time scale s_n in (i).

In order to formulate our results we need some more ingredients:

1. Let $(Y_t)_{t\ge 0}$ be a standard Fisher–Wright diffusion on $[0,1]$, that is, the solution of

$$dY_t = \sqrt{Y_t(1-Y_t)}\, dW_t \tag{1.13}$$

(W_t is a standard Brownian motion) and let $Q_t(\cdot,\cdot)$ be its transition semigroup. It is known that 0 and 1 are accessible boundary points for Y_t [see, e.g., Ethier and Kurtz (1986), Proposition 10.2.8]. Hence $\lim_{t\to\infty} \mathbf{P}^\theta[Y_t = 1] = 1 - \lim_{t\to\infty} \mathbf{P}^\theta[Y_t = 0] = \theta$.

2. It turns out that there are two main regimes of clustering. For their classification we will need the recurrent potential kernel of the random walk induced by a:

$$A(\zeta,\xi) = \sum_{m=0}^{\infty} \big(a^{(m)}(\zeta,\zeta) - a^{(m)}(\zeta,\xi)\big). \tag{1.14}$$

Furthermore, let $A(n) = \sup_{\xi\in\Xi_n} A(0,\xi) = A(0,\zeta)$ for any ζ with $\|\zeta\| = n$. As usual, $a^{(m)}$ denotes the m-step transition probability induced by a. The existence of the recurrent potential kernel is assured, for example, by Kemeny, Snell and Knapp [(1976), Corollary 9-29]. Note that an irreducible recurrent random walk on an infinite denumerable Abelian group is null recurrent. [For random walks on $\mathbb{Z}^d$ the existence is due to Spitzer (1964), P12.1 and P28.4.]

The kernel a is called *critical* (or *critically recurrent*) if it is recurrent and

$$\log(k)\big[\log(r_k N^k) - \log(r_{k+1}N^{k+1})\big] \tag{1.15}$$

is bounded. For example, the geometrical kernel a_1 is critical. On the other hand, the recurrent kernels a_c with $c > 1$ are called *strongly recurrent*.

3. In the case a critical and for $\alpha \in [0, 1]$ let the α-space scale be a function f_α: $\mathbb{N}_0 \to \mathbb{N}_0$ (depending only on the potential kernel) that is chosen such that

$$\alpha = \lim_{n\to\infty} \frac{A(f_\alpha(n))}{A(n)} \tag{1.16}$$

and let the time scale be $s_n = N^n A(n)$.

THEOREM 1 (Cluster formations in the case a critical). *Suppose that* (1.15) *and* (1.16) *hold. Then*

(a) $\mathscr{L}^\mu\big(\Theta_{f_\alpha(n(t))}(\mathbb{X}(t))\big) \Rightarrow \mathscr{L}^\theta(Y_{\hat\alpha})$ *as* $t \to \infty$,

(b) $\mathscr{L}^\mu\big({}^{f_\alpha}(\mathbb{X}(t))\big) \Rightarrow \nu_\theta(\hat\alpha) := \int Q_{\hat\alpha}(\theta, d\rho)\pi_\rho$ *as* $t \to \infty$,

where $\mu \in \mathscr{M}_\theta$, $\hat\alpha := -\log\alpha$ *and* π_ρ *is the product measure concentrated on* $\{0, 1\}^\Xi$ *with intensity* $\rho = \langle \pi_\rho, x_0\rangle$.

REMARKS. (i) Theorem 1 states that for fixed α there exists one possible limit field $\nu(\hat\alpha)$ independent of the particular choice of the (critical) a and $\mathbb{X}$ converges to $\nu(\hat\alpha)$ when rescaled by f_α and (s_n). The asymptotic behavior of f_α thus measures the speed at which clusters grow. There are mainly (i.e., with some additional monotonicity conditions) three sizes of clusters:

$$\begin{array}{lll}
\text{small clusters} & \text{when } \dfrac{f_\alpha(n)}{n} \to 0 & \text{as } n \to \infty, \text{ for } \alpha < 1,\\[2ex]
\text{medium clusters} & \text{when } \lim\limits_{n\to\infty} \dfrac{f_\alpha(n)}{n} \in]0, 1[& \text{for } \alpha \in]0, 1[,\\[2ex]
\text{large clusters} & \text{when } \dfrac{f_\alpha(n)}{n} \to 1 & \text{as } n \to \infty, \text{ for } \alpha > 0.
\end{array}$$

For instance the above cases can occur if we choose

$$\begin{array}{lll}
r_k = \vartheta k N^{-k} & \text{and} & f_\alpha(n) = n^\alpha,\\
r_k = \vartheta N^{-k} & \text{and} & f_\alpha(n) = \alpha n,\\
r_k = \vartheta k^{-\log k} N^{-k} & \text{and} & f_\alpha(n) = n\left(1 + \dfrac{\log\alpha}{2\log n}\right)
\end{array}$$

(for some normalizing constants ϑ).

(ii) We can choose $f_0 \equiv 0$. Hence (1.7) is included in (b), since $\nu(\hat{0}) = \nu(\infty) = (1-\theta)\delta_{\mathbf{0}} + \theta\delta_{\mathbf{1}}$.

(iii) Note that the statements of Theorem 1 do not depend on the choice of g. This is also true for the Theorems 2, 4 and 5. The asymptotic behavior of $\mathbb{X}(t)$ as $t \to \infty$ is determined by the interaction kernel rather than by the diffusion coefficient. For a detailed discussion of this point, see Cox, Fleischmann and Greven (1996).

Let us now turn to the case a strongly recurrent. Here the picture is by far not as complete as in the case a critical. In fact, a statement such as Theorem 1(b) cannot be expected. This case is the analogue of the $d = 1$ case for finite variance interaction kernels on $\mathbb{Z}^d$. Despite this we cannot expect an invariance principle such as Arratia's (1982) for the voter model on $\mathbb{Z}$. In fact this greatly depends on the linear structure of $\mathbb{Z}$ and on the comparably simple structure of the voter model. Conceptually Arratia's work is based on nearest neighbor interaction. Recently extensions have been made to Arratia's result which are concerned with stochastic partial differential equation models [Tribe (1993), Section 7] or more general interaction kernels in the voter model [Cox and Durrett (1995), Theorem 4], but these still rely on the linear structure. In order to circumvent this problem we use the idea of renormalization via block averages [recall (1.12)] and establish that the limiting density chain $(Z_m^{N,t}) = \text{w-lim}_{n\to\infty} \Theta_{n-m}(\mathbb{X}(ts_n))$ (with m as time parameter) exists. In fact, the distribution of the limiting chain can be determined. Namely, the moments can be expressed in terms of a coalescing system with motion given by weak limits $\gamma(t)$ of rescaled random walks on Ξ. This is done in Section 5.

In order to bring some more light into the structure of $(Z_m^{N,t})$ we then let $N \to \infty$ to obtain even a Markov chain. To describe the transition probabilities of this chain we need the following diffusion X_t^θ on $[0,1]$ given by $X_0^\theta = \theta$ and

$$dX_t^\theta = \sqrt{2(c-1)X_t^\theta(1-X_t^\theta)}\,dW_t + (\theta - X_t^\theta)\,dt. \tag{1.17}$$

Then $\mathscr{L}\left[X_t^\theta\right]$ converges weakly to the unique invariant law of (1.17) as $t \to \infty$, which is known to be the β-distribution

$$\mathscr{L}\left[X_\infty^\theta\right] = B\left(\frac{1}{c-1}\theta,\ \frac{1}{c-1}(1-\theta)\right) \tag{1.18}$$

[see, e.g., Ethier and Kurtz (1986), Chapter 10, Lemma 2.1].

Assume a_c is strongly recurrent ($c > 1$). Here again $N^n A(n)$ would give the right time scale, but since $A(n)$ can be computed to be $\kappa(N)c^n$ with $\kappa(N)(Nc)^{-1} \to 1$ as $N \to \infty$, we prefer to let

$$s_n = (Nc)^{n+1}. \tag{1.19}$$

THEOREM 2 (Cluster formations in the case a strongly recurrent). (a) *For any N and $t>0$ there exists a nonnegative martingale $(Z_m^{N,t})_{m\in\mathbb{Z}}$ such that*

$$\mathscr{L}^{\mu}\left[(\Theta_{n-m}(\mathbb{X}(ts_n)))_{m\in\mathbb{Z}}\right] \Rightarrow \mathscr{L}^{\theta}\left[(Z_m^{N,t})_{m\in\mathbb{Z}}\right] \quad \text{as } n\to\infty,$$

where $\mu\in\mathscr{M}_\theta$. This martingale has the following properties:

$$\text{(b)}\qquad \mathscr{L}^{\theta}(Z_m^{N,t}) \Rightarrow \theta\delta_1+(1-\theta)\delta_0 \quad \text{as } m\to\infty,$$

$$\mathscr{L}^{\theta}(Z_m^{N,t}) \Rightarrow \delta_\theta \quad \text{as } m\to-\infty;$$

$$\text{(c)}\qquad (Z_m^t)_{m\in\mathbb{Z}} := w\text{-}\lim_{N\to\infty}\left[(Z_m^{N,t})_{m\in\mathbb{Z}}\right] \text{ exists and is Markov.}$$

The transition mechanism of (Z_m^t) is given by

$$\mathscr{L}\left[Z_m^t \middle| Z_{m-1}^t=\rho\right] = \begin{cases} \delta_\rho, & m<0,\\ \mathscr{L}[X_t^\rho], & m=0,\\ \mathscr{L}[X_\infty^\rho], & m>0. \end{cases} \tag{1.20}$$

REMARKS. (i) At first glance the appearance of X_t^θ in Theorem 2 might be surprising. The key for understanding its meaning is the duality (Lemma 5.5) of X_t^θ to the so-called death-escape process. This is a modification of the pure death process (Definition 3.1) which is known to be dual to the Fisher–Wright diffusion with no drift.

(ii) The process $(Z_m^{N,t})$ is *not* Markov for fixed N since the influence of Θ_{n-m+2} on Θ_{n-m} given Θ_{n-m+1} does not vanish as $n\to\infty$. However, computer simulations show that $(Z_m^{N,t})$ is even for small N not "too far off" from the limiting structure $N\to\infty$.

(iii) Dawson and Greven (1993b) obtain their "interaction chain" by letting $N\to\infty$ for fixed n. A simple computation shows that letting $n\to\infty$ for that chain and rescaling time properly yields the same chain (Z_m^t). Thus the order of the limits can be interchanged. To see that the stable laws there approximate our $\mathscr{L}[X_\infty^\theta]$ one needs Baillon, Clément, Greven and den Hollander [(1995), Theorem 1(a)].

Theorem 2 asserts in particular that clusters grow all at "maximum speed." Note the difference between large clusters in the case a critically recurrent and clusters in the case, a strongly recurrent. In the former case, $f_\alpha(n)-n\to-\infty$ as $n\to\infty$ for $\alpha<1$; thus Theorem 2(b) would not hold.

Finite systems versus infinite systems. Since all computers known to the author so far (May 11, 1995) are of finite size, simulations have to be restricted to finite versions of the model. On the other hand, finite systems can be considered in their own right. They model a finite nature and the infinite system can be regarded as an idealization for analytical convenience only. So the questions arise: How well do finite systems approximate the infinite system (and vice versa)? How long can a finite system be observed until it "feels" its finiteness and which effects of finiteness do occur?

A number of approaches have been used in the literature for various models [see, e.g., Durrett and Schonmann (1988) or Dawson and Gärtner (1988)]. We will proceed in the fashion of the finite systems scheme suggested by Cox and Greven (1990, 1994b): The system is dominated by the macroscopic variable of the block averages. Roughly speaking it relaxes to an "equilibrium state" with intensity θ, given that the block average is θ. This relaxation takes place faster than the fluctuation of the block averages. In the case a transient these equilibria are the invariant measures ν_θ, while in the case a critical we have to take the $\nu_\theta(\hat{\alpha})$ (introduced in Theorem 1) instead. In the case a strongly recurrent, however, the finite systems scheme does not work. This is connected with the fact that the intensity, that is the block averages of components, alone does not characterize the system above any more. Hence the (macroscopic) associated process $(\widetilde{Z}_m^{N,t})$ is *not* Markov.

We first define the finite system $\mathbb{X}_n(t)$ and (in case of criticality) the scaled finite system ${}^f\mathbb{X}_n(t)$ as the solution of the restricted SSDE:

$$dx_{n,\xi}(t) = \left(\sum_{\zeta\in\Xi_n} a_n(\xi,\zeta)(x_{n,\zeta}(t) - x_{n,\xi}(t))\right)dt \tag{1.21}$$
$$+\sqrt{g(x_{n,\xi}(t))}\,dW_\xi(t) \qquad (\xi\in\Xi_n),$$

where

$$a_n(\xi,\zeta) = \sum_{\substack{\zeta'\in\Xi\\ \zeta'\equiv\zeta(\mathrm{mod}\,\Xi_n)}} a(\xi,\zeta') \tag{1.22}$$

and

$${}^f x_{n,\xi}(t) = x_\zeta(ts_n), \quad \text{where } \zeta = S_{f(n)}^{-1}\xi, \tag{1.23}$$

$$\mathscr{L}^\mu(\mathbb{X}_n(0)) = \mu_n := \mu|_{\Xi_n}. \tag{1.24}$$

Note that the space scale here does *not* depend on t as before, but on the finite system size n.

By speeding up time by the factor s_n we expect the intensity $\Theta_n(\mathbb{X}_n(ts_n))$ to start to fluctuate and to tend to some nontrivial process $\widetilde{Y}_t$. We even hope that $\mathbb{X}_n(t)$ [resp. ${}^{f_\alpha}\mathbb{X}_n(t)$] "relaxes" fast enough, so its limiting distribution given $\widetilde{Y}_t = \rho$ is ν_ρ [resp. $\nu_\rho(\hat{\alpha})$]. In fact, an integral statement of this heuristic holds in the cases a transient or critical, where $\widetilde{Y}_t$ turns out to be a Fisher–Wright diffusion running at double speed.

In the case a transient a prominent role is played by the Green function

$$G(\xi,\zeta) = \sum_{m=0}^{\infty} a^{(m)}(\xi,\zeta). \tag{1.25}$$

Its role is analogous to that of the recurrent potential kernel for the case a critically recurrent.

Assume a to be transient, $g(x) = x(1-x)$ and let $(\nu_\theta | \theta \in [0,1])$ be the family of invariant measures. Let $G = G(0,0)$ and

$$V := \mathbf{E}^0\left[\exp\left(-\frac{1}{2}\int_0^\infty \mathbb{1}_{\{X_s=0\}}\, ds\right)\right], \tag{1.26}$$

where $(X_s)_{s\geq 0}$ is the continuous time random walk associated with $a(\cdot,\cdot)$ (see Section 2.1).

Let the time scale be

$$s_n = \frac{G}{1-V} N^n. \tag{1.27}$$

To put the latter discussion in perspective we give the following result for the transient case.

THEOREM 3 (Finite system, case a transient). *Under these assumptions for* $t > 0$ *the following situations hold*:

(a) $$\mathscr{L}^\mu(\Theta_n(\mathbb{X}_n(ts_n))) \Rightarrow \mathscr{L}^\theta(Y_{2t}) \quad \textit{as } n \to \infty,$$

(b) $$\mathscr{L}^\mu(\mathbb{X}_n(ts_n)) \Rightarrow \int Q_{2t}(\theta, d\rho)\nu_\rho \qquad \textit{as } n \to \infty,$$

where $\mu \in \mathscr{M}_\theta$.

REMARKS. (i) The condition on g can be dropped, but then $\widetilde{Y}_t$ [the limiting process of $\Theta_n(\mathbb{X}_n(ts_n))$] does not have such a simple form. We do not stress this point here. In the lattice case a stronger version of Theorem 3 can be found in Cox, Greven and Shiga [(1994), Theorem 2].

(ii) In the voter model a similar statement holds, when s_n is replaced by GN^n. For the lattice case of this, see Cox [(1989), Theorem 2 and 3]. For the case a critical, see Cox and Greven [(1991), Theorem 1].

Assume now a to be critical. Again things happen to depend only on the recurrent potential kernel.

THEOREM 4 (Finite system, case a critical). *Let* α, f_α *and* $s_n = N^n A(n)$ *be as in Theorem* 1. *Then for* $t > 0$ *the following situations hold*:

(a) $$\mathscr{L}^\mu\big(\Theta_{f_\alpha(n)}(\mathbb{X}_n(ts_n))\big) \Rightarrow \mathscr{L}^\theta(Y_{2t+\hat{\alpha}}) \quad \textit{as } n \to \infty,$$

(b) $$\mathscr{L}^\mu\big({}^{f_\alpha}\mathbb{X}_n(t)\big) \Rightarrow \int Q_{2t}(\theta, d\rho)\nu_\rho(\hat{\alpha}) = \int Q_{2t+\hat{\alpha}}(\theta, d\rho)\pi_\rho \quad \textit{as } n \to \infty,$$

where $\mu \in \mathscr{M}_\theta$.

Let a_c be strongly recurrent. Considerably less can be said in this situation since Theorem 2 is weaker than Theorem 1. Again we use the slightly modified time scale

$$s_n = (Nc)^{n+1}.$$

THEOREM 5 (Finite system, case a strongly recurrent). (a) *For any N and $t > 0$ there is a nontrivial martingale $(\widetilde{Z}_m^{N,t})_{m=0,1,\dots}$ such that*

$$\mathscr{L}^{\mu}\big[(\Theta_{n-m}(\mathbb{X}_n(ts_n)))_{m=0,1,\dots}\big] \Rightarrow \mathscr{L}^{\theta}\big[(\widetilde{Z}_m^{N,t})_{m=0,1,\dots}\big] \quad \text{as } n \to \infty,$$

where $\mu \in \mathscr{M}_\theta$. This martingale has the following properties:

(b) $$\mathscr{L}^{\theta}\big(\widetilde{Z}_m^{N,t}\big) \Rightarrow \theta\delta_1 + (1-\theta)\delta_0 \quad \text{as } m \to \infty;$$

(c) $$(\widetilde{Z}_m^t)_{m=0,1,\dots} := w\text{-}\lim_{N\to\infty}\big[(\widetilde{Z}_m^{N,t})_{m=0,1,\dots}\big] \text{ exists and is Markov.}$$

The transition mechanism of $(\widetilde{Z}_m^t)$ is given by

$$\mathscr{L}\big[\widetilde{Z}_m^t \big| \widetilde{Z}_{m-1}^t = \rho\big] = \begin{cases} \mathscr{L}[Y_{2t}^{\rho}], & m = 0, \\ \mathscr{L}[X_\infty^{\rho}], & m > 0. \end{cases} \tag{1.28}$$

REMARK. Compare $(\widetilde{Z}_m^t)$ with (Z_m^t). The transition mechanisms coincide except for $m = 0$. Here the difference between the infinite and the finite system becomes clear. In the infinite system there are blocks at level $m = -1$ with deterministic intensity θ that puts a drift on the fluctuation of $(Z_0^t)_{t\geq 0}$ while in the finite system these bigger blocks do not exist and thus the drift is missing.

Outline. The rest of the paper is organized as follows. Since the system considered is for $g(x) = x(1-x)$ in duality with delayed coalescing random walks, we develop in Section 2 some first hitting time asymptotics for random walks with scaled initial points on a rather general class of Abelian groups by using the Green function and recurrent potential properties. These properties will be used in the investigation of systems of coalescing random walks in Section 3. In Section 4 we do moment calculations in our original problem via a duality relation in the special case $g(x) = x(1-x)$. Based on this, generalizations will be obtained by coupling and comparison arguments. This will suffice to give the proofs of Theorems 1, 3 and 4. Since Theorems 2 and 5 are somewhat different, their proofs are deferred to Section 5.

2. Random walk estimates. The goal of this section is to derive results on the asymptotic behavior of hitting times of 0 for sequences of initial points which typically move away from 0. The key result is Proposition 2.7 in Section 2.4.4.

2.1. *Preparations.* First we develop some more general results on random walks on a countably infinite Abelian group $(\Lambda, +)$ and then give examples in $\mathbb{Z}^d$ and Ξ.

Let (G_n) be a sequence of subgroups of Λ. Assume that we can choose for any $n \in \mathbb{N}$ a complete system $\Lambda_n \subset \Lambda$ of representatives for the quotient group Λ/G_n such that $\Lambda_1 \subset \Lambda_2 \subset \cdots$ and $\lim_{n\to\infty} \Lambda_n = \Lambda$. For example, think of

$\Lambda = \mathbb{Z}^d$, $G_n = n\mathbb{Z}^d$ and $\Lambda_n = (]-n/2, n/2] \cap \mathbb{Z})^d$. Further let $p(\cdot, \cdot)$ be the transition kernel of an irreducible random walk on Λ. Let $p_n(\cdot, \cdot)$ be the kernel of the induced random walk on Λ_n. That is, $p_n(x, y) = \sum_{g \in G_n} p(x, y+g)$. Let $(X(t))_{t\geq 0}$ [resp. $(X_n(t))_{t\geq 0}$] denote the induced continuous time random walks, that is, with transition probabilities

$$p(t; x, y) := \mathbf{P}(X(t) = y | X(0) = x) = e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} p^{(k)}(x, y), \tag{2.1}$$

$$\begin{aligned} p_n(t; x, y) &:= \mathbf{P}(X_n(t) = y | X(0) = x) = e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} p_n^{(k)}(x, y) \\ &= e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{g \in G_n} p^{(k)}(x, y+g). \end{aligned} \tag{2.2}$$

The key role is played by the recurrent potential kernel [recall (1.14)]

$$A(x, y) = \sum_{m=0}^{\infty} \big(p^{(m)}(x, x) - p^{(m)}(x, y)\big) \tag{2.3}$$

which is well defined for either recurrent or transient random walk. In the latter case we have, in addition,

$$A(x, y) = G(x, x) - G(x, y) = G - G(x, y), \tag{2.4}$$

where $G(x, y) = \sum_{m=0}^{\infty} p^{(m)}(x, y)$ and $G = G(0, 0)$. Further let

$$A(n) = \sup_{x \in \Lambda_n} A(0, x) \tag{2.5}$$

and for later technical convenience let (a_n) be a sequence such that

$$\lim_{n\to\infty} \frac{a_n}{A(n)} = 1. \tag{2.6}$$

The purpose of this section is the investigation of the first hitting times of the origin

$$\tau = \inf\{t \geq 0 | X(t) = 0\}, \tag{2.7}$$

$$\tau_n = \inf\{t \geq 0 | X_n(t) = 0\}. \tag{2.8}$$

Since the random walks will typically be started from initial points (x_n) far away, we shall consider τ and τ_n but scaled by s_n. Here

$$s_n = a_n |\Lambda_n|. \tag{2.9}$$

We have to make some more assumptions on the random walk.

DEFINITION 2.1 (Diffusive random walk). The random walk $X(t)$ [and its kernel $p(\cdot,\cdot)$] is called *diffusive* if the following assumptions hold:

$$\exists\, K<\infty:\ \sup_{\substack{m\geq 0,\, n\geq 0\\ x\in\Lambda_n}} \big(p_n^{(m)}(0,x)-p^{(m)}(0,x)\big)|\Lambda_n|<K, \tag{2.10}$$

$$\sup_{x\in\Lambda_n}\big||\Lambda_n|p_n^{([ts_n])}(0,x)-1\big|\to 0 \quad \text{as } n\to\infty,\ \forall\, t>0. \tag{2.11}$$

There exists a sequence $(c_n)\ll(a_n)$ such that

$$|\Lambda_n|\sup_{m\geq c_n|\Lambda_n|t} p^{(m)}(0,0)\to 0 \quad \text{as } n\to\infty, \forall\, t>0, \tag{2.12}$$

$$\frac{1}{a_n}\sum_{m=0}^{c_n|\Lambda_n|} p^{(m)}(0,0)\to 1 \quad \text{as } n\to\infty, \tag{2.13}$$

$$\frac{1}{a_n}\sup_{x\in\Lambda_n}\left|\sum_{m=c_n|\Lambda_n|}^{\infty}\big[p^{(m)}(0,0)-p^{(m)}(0,x)\big]\right|\to 0 \quad \text{as } n\to\infty. \tag{2.14}$$

Here we used the notation $(c_n)\ll(a_n)$ for $c_n/a_n\to 0$ as $n\to\infty$.

2.2. *Scaled limits of hitting times.* Assume $X(t)$ to be diffusive (either transient or recurrent) and let $(x_n)_{n\in\mathbb{N}}$ be a sequence with $x_n\in\Lambda_n$, $n\in\mathbb{N}$, such that

$$\alpha:=\lim_{n\to\infty}\frac{A(0,x_n)}{A(n)} \tag{2.15}$$

exists. Denote by $\mathscr{E}(\mu)$ the exponential distribution with mean μ. By $\mathscr{L}^x$ ($\mathbf{P}^x$, $\mathbf{E}^x$) we denote the law (probability, expectation) with respect to the initial point x. By δ_∞ we denote the unit mass at $+\infty\in\mathbb{R}\cup\{-\infty,+\infty\}$. That is, $\mathbf{P}[X>x]=1$ $\forall x\in\mathbb{R}$ if $\mathscr{L}[X]=\delta_\infty$.

PROPOSITION 2.2 (Diffusive random walk on Λ). *We have*

(i) $$\mathscr{L}^{x_n}\left(\frac{\tau}{s_n}\right)\Rightarrow(1-\alpha)\delta_0+\alpha\delta_\infty \quad \textit{as } n\to\infty;$$

(ii) $$\mathscr{L}^{x_n}\left(\frac{\tau_n}{s_n}\right)\Rightarrow(1-\alpha)\delta_0+\alpha\cdot\mathscr{E}(1) \quad \textit{as } n\to\infty.$$

PROOF. It is enough to show the convergence of the Laplace transforms $T_n(\lambda)=\mathbf{E}^{x_n}[\exp(-\lambda\tau/s_n)]$ and $T'_n(\lambda)=\mathbf{E}^{x_n}[\exp(-\lambda\tau_n/s_n)]$. We will show

$$T_n(\lambda)\to 1-\alpha \quad \text{as } n\to\infty,$$
$$T'_n(\lambda)\to 1-\alpha+\frac{\alpha}{1+\lambda} \quad \text{as } n\to\infty.$$

By a simple first hitting time decomposition we obtain

$$T'_n(\lambda) \sim \frac{\sum_{m=0}^{\infty} p_n^{(m)}(0, x_n)\exp(-\lambda m/s_n)}{\sum_{m=0}^{\infty} p_n^{(m)}(0,0)\exp(-\lambda m/s_n)} \quad \text{as } n \to \infty. \tag{2.16}$$

We multiply by $1/a_n$ and split the numerator in three parts

$$\begin{aligned} &\frac{1}{a_n}\sum_{m=0}^{\infty}\big[p_n^{(m)}(0,x_n) - p^{(m)}(0,x_n)\big]\exp\bigg(\frac{-\lambda m}{s_n}\bigg) \\ &\quad + \frac{1}{a_n}\sum_{m=0}^{\infty}\big[p^{(m)}(0,x_n) - p^{(m)}(0,0)\big]\exp\bigg(\frac{-\lambda m}{s_n}\bigg) \\ &\quad + \frac{1}{a_n}\sum_{m=0}^{\infty} p^{(m)}(0,0)\exp\bigg(\frac{-\lambda m}{s_n}\bigg). \end{aligned} \tag{2.17}$$

The three sums are now estimated separately:

(i) $$\begin{aligned} &\lim_{n\to\infty}\sup_{x\in\Lambda_n}\bigg|\frac{1}{a_n}\sum_{m=0}^{\infty}\big[p_n^{(m)}(0,x) - p^{(m)}(0,x)\big]\exp\bigg(\frac{-\lambda m}{s_n}\bigg) - \frac{1}{\lambda}\bigg| \\ &\quad = \lim_{n\to\infty}\sup_{x\in\Lambda_n}\bigg||\Lambda_n|\int_0^{\infty}\big[p_n^{([ts_n])}(0,x) - p^{([ts_n])}(0,x)\big]\exp(-\lambda t)\,dt - \frac{1}{\lambda}\bigg| = 0 \end{aligned}$$

since the integrand is bounded by $(K/|\Lambda_n|)e^{-\lambda t}$ and is of order $(1/|\Lambda_n|)e^{-\lambda t}$ [by (2.10) and (2.11)];

(ii) $$\begin{aligned} &\lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{\infty}\big[p^{(m)}(0,0) - p^{(m)}(0,x_n)\big]\exp\bigg(\frac{-\lambda m}{s_n}\bigg) \\ &\quad = \lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{c_n|\Lambda_n|}\big[p^{(m)}(0,0) - p^{(m)}(0,x_n)\big] \quad \text{[by (2.14)]} \\ &\quad = \lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{\infty}\big[p^{(m)}(0,0) - p^{(m)}(0,x_n)\big] = \alpha \quad \text{[by (2.14)]}; \end{aligned}$$

(iii) $$\begin{aligned} &\lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{\infty} p^{(m)}(0,0)\exp\bigg(\frac{-\lambda m}{s_n}\bigg) \\ &\quad = \lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{c_n|\Lambda_n|} p^{(m)}(0,0)\exp\bigg(\frac{-\lambda m}{s_n}\bigg) \quad \text{[by (2.12)]} \\ &\quad = 1 \quad \text{[by (2.13)]}. \end{aligned}$$

Putting the pieces together we obtain the convergence of the numerator to $1/\lambda - \alpha + 1$. A similar expansion shows that the denominator converges to $1/\lambda + 1$. So we are done with the finite case. For the infinite case note that the first term of the expansion vanishes. So the convergence of the Laplace transform is obtained the same way. □

Now look deeper into the case $X(t)$ transient. Here we can choose $a_n \equiv G$ and (2.13) and (2.14) trivially hold with any sequence $c_n \gg |\Lambda_n|^{-1}$.

So assume $X(t)$ to be transient and diffusive. Let

$$s_n = G|\Lambda_n|. \tag{2.18}$$

Assume that

$$\gamma = \lim_{n\to\infty} G(0, x_n) \tag{2.19}$$

exists.

COROLLARY 2.3 (Transient diffusive random walk on Λ). *Under these assumptions*,

(i) $$\mathscr{L}^{x_n}\left(\frac{\tau}{s_n}\right) \Rightarrow \frac{\gamma}{G}\delta_0 + \left(1 - \frac{\gamma}{G}\right)\delta_\infty \quad \text{as } n \to \infty,$$

(ii) $$\mathscr{L}^{x_n}\left(\frac{\tau_n}{s_n}\right) \Rightarrow \frac{\gamma}{G}\delta_0 + \left(1 - \frac{\gamma}{G}\right)\mathscr{E}(1) \quad \text{as } n \to \infty.$$

2.3. *Application to* $\mathbb{Z}^d$. As a first example we give a well-known result on symmetric Bernoulli random walk on $\mathbb{Z}^d$.

Let $\Lambda_n =]-n/2, n/2]^d \cap \mathbb{Z}^d$, (b_n) be some real sequence $n/2 > b_n \uparrow \infty$ and

$$s_n = \begin{cases} \dfrac{2}{\pi} n^2 \log n, & \text{if } d = 2, \\ Gn^d, & \text{if } d \geq 3. \end{cases} \tag{2.20}$$

PROPOSITION 2.4. (a) *If $d \geq 3$, then uniformly in all sequences $(x_n)_{n\in\mathbb{N}}$ with $x_n \in \Lambda_n$, $n \in \mathbb{N}$, and such that $|x_n| > b_n$*,

$$\mathbf{P}^{x_n}(\tau_n/s_n > t) \to e^{-t} \quad \text{as } n \to \infty.$$

(b) *If $d = 2$, let $\alpha \in [0, 1]$ and assume $|x_n| \sim n^\alpha$. Then*

$$\mathbf{P}^{x_n}(\tau_n/s_n > t) \to \alpha e^{-t} \quad \text{as } n \to \infty,$$

$$\mathbf{P}^{x_n}(\tau/s_n > t) \to \alpha \quad \text{as } n \to \infty.$$

REMARKS. 1. Part (a) is Theorem 4 of Cox (1989), while (b) is a combination of this and a result of Erdös and Taylor [(1960), equation (2.16)].

2. The Bernoulli random walk in $\mathbb{Z}^1$ is not diffusive. Indeed $A(0, x) = |x|$ [see Spitzer (1964), E29.1] is not slowly varying.

PROOF OF PROPOSITION 2.4. Since $|\Lambda_n| = n^d$ we can choose $a_n = (2/\pi) \times \log n$ if $d = 2$ [see P12.3 of Spitzer (1964)]. It remains to verify diffusiveness.

Since there exists a $K < \infty$ such that

$$p^{(m)}(0, x) \leq Km^{-d/2} \exp\left(-\frac{d|x|^2}{2m}\right) \tag{2.21}$$

[see, e.g., P7.10, Spitzer (1964)], one easily derives (2.10). Then (2.11) is implied by Proposition 2.8 of Cox (1989), which is obtained by a Bhattacharya–Rao expansion. By (2.21),

$$mp^{(m)}(0,x) \le Km^{1-d/2} \to 0 \quad \text{as } m \to \infty$$

if $d \ge 3$. This implies (2.12).

Assume now $d = 2$. Let $c_n = \sqrt{\log n}$. Then (2.12) follows from (2.21). Since $p^{(m)}(0,0) \sim (1/\pi)1/m$ we have

$$\sum_{m=0}^{n^2\sqrt{\log n}} p^{(m)}(0,0) \sim \frac{1}{\pi}\log\Big(n^2\sqrt{\log n}\Big) \sim \frac{2}{\pi}\log n, \tag{2.22}$$

so (2.13) is valid. Again by (2.21),

$$\big|p^{(m)}(0,0) - p^{(m)}(0,x)\big| \le K\frac{1}{m}\left(1 - \exp\left(\frac{-|x|^2}{m}\right)\right) \qquad \forall\, x, m, \tag{2.23}$$

so

$$\sum_{m=M}^{\infty} \big|p^{(m)}(0,0) - p^{(m)}(0,x)\big| \le \frac{2K|x|^2}{M}. \tag{2.24}$$

Putting $M = n^2\sqrt{\log n}$ yields (2.14). □

2.4. *Application to* Ξ. In order to apply Proposition 2.2 and Corollary 2.3 to random walks on Ξ we have to calculate the m-step transition probabilities $p^{(m)}$. This is a relatively simple task due to the special geometry of Ξ. We then compute the potential kernels and verify the diffusiveness assumptions for the cases $X(t)$ transient and critical separately.

2.4.1. *Computation of the transition probabilities.* Introduce

$$\hat{\Xi} := \{(a_k)_{k\in\mathbb{N}}\colon\ a_k \in \{0,\dots,N-1\}\}$$

with addition componentwise modulo N and the scalar product

$$\langle a, \xi\rangle = \exp\left(\frac{2\pi i}{N}\sum_{k=1}^{\infty} a_k\xi_k\right).$$

Hence $\hat{\Xi}$ is the character group of Ξ. Now some Fourier transformations yield the desired transition probabilities [see Fleischmann and Greven (1994), Section 2a].

For $k = 1, 2, \dots$ let $f_k = r_0 + \cdots + r_{k-1} - (1/(N-1))r_k$. [Recall from (1.5) that $a(\xi,\zeta) = r_k/R_k$ for $\|\zeta - \xi\| = k$.] Then

$$p^{(m)}(0,\xi) = (N-1)\sum_{k>\|\xi\|} N^{-k}(f_k)^m + (\mathbb{1}_{\{0\}}(\xi) - 1)N^{-\|\xi\|}(f_{\|\xi\|})^m, \tag{2.25}$$

$$\begin{aligned} p(t;0,\xi) = (N-1)&\sum_{k>\|\xi\|} N^{-k}\exp(-t(1-f_k)) \\ &+ (\mathbb{1}_{\{0\}}(\xi) - 1)N^{-\|\xi\|}\exp(-t(1-f_{\|\xi\|})). \end{aligned} \tag{2.26}$$

Write also $p^{(m)}(n)$ for $p^{(m)}(0,\xi)$ with $\|\xi\| = n$. By restricting the random walk to $\Lambda_n := \Xi_n$ (note $|\Xi_n| = N^n$) the r_k transform to

$$r_{n,k} = \begin{cases} r_k\left(1 - \sum_{l=n+1}^{\infty} r_l\right)^{-1}, & \text{if } k \le n, \\ 0, & \text{otherwise.} \end{cases} \tag{2.27}$$

Hence we put $f_{n,k} = r_{n,0} + \cdots + r_{n,k-1} - r_{n,k}/(N-1)$ to obtain from (2.25) and (2.26) the transition probabilities in the finite setting:

$$\begin{aligned} p_n^{(m)}(0,\xi) = (N-1) &\sum_{k=\|\xi\|+1}^{n} N^{-k}(f_{n,k})^m \\ &+ (\mathbb{1}_{\{0\}}(\xi) - 1)N^{-\|\xi\|}(f_{n,\|\xi\|})^m + N^{-n}, \end{aligned} \tag{2.28}$$

$$\begin{aligned} p_n(t;0,\xi) = (N-1) &\sum_{k=\|\xi\|+1}^{n} N^{-k}\exp\{-t(1-f_{n,k})\} \\ &+ (\mathbb{1}_{\{0\}}(\xi) - 1)N^{-\|\xi\|}\exp\{-t(1-f_{n,\|\xi\|})\} + N^{-n}. \end{aligned} \tag{2.29}$$

Note that (2.10) is always valid by symmetry.

2.4.2. *Case $X(t)$ transient.* Now look into the case $X(t)$ transient in detail.

LEMMA 2.5 (Transient random walk on Ξ). *A transient random walk on Ξ is diffusive in the sense of Definition* 2.1.

PROOF. The quantity G can be explicitly expressed in terms of the f_k. By (2.25), G equals

$$G = (N-1)\sum_{k=1}^{\infty} \frac{N^{-k}}{1-f_k}.$$

By transience $G < \infty$ and hence

$$\liminf_{n\to\infty} N^n \sum_{k=n}^{\infty} r_k = \infty. \tag{2.30}$$

In particular, for $c > 1/N$, let be $G_c(\cdot,\cdot)$ the Green function associated with the geometrical kernel a_c. Let $G_c = G_c(0,0)$. Then

$$G_c = \frac{Nc(N-1)^2}{N^2c-1}\sum_{k=1}^{\infty} c^k. \tag{2.31}$$

Thus $G_c < \infty$ iff $c < 1$. In this case,

$$G_c = \frac{Nc^2(N-1)^2}{(1-c)(N^2c-1)}. \tag{2.32}$$

Let T_n denote the first exit time of Ξ_n:

$$T_n := \inf\{t \ge 0\colon\ X(t) \in \Xi \setminus \Xi_n\}. \tag{2.33}$$

The $\mathscr{L}^{\xi}(T_n)$ coincide for all $\xi \in \Xi_n$. Hence by the Markov property, $\mathscr{L}^{\xi}(T_n) = \mathscr{E}(\mu)$ for some $\mu \geq 0$ [recall $\mathscr{E}(\mu)$ is exponential with mean μ]. Note that μ does not change if we replace r_{n+1} by $r'_{n+1} = \sum_{k=n+1}^{\infty} r_k$ and r_k by 0 for $k > n+1$. Denote the corresponding transition probabilities by p'. Then by (2.26) for $t \to 0$,

$$\begin{aligned} \sum_{\zeta \in \Xi \setminus \Xi_n} p(t, 0, \zeta) &= \sum_{\zeta \in \Xi_{n+1} \setminus \Xi_n} p'(t, 0, \zeta) \\ &= \frac{N-1}{N} \left(1 - \exp\left\{ -t \frac{N}{N-1} r'_{n+1} \right\} \right) \\ &= t r'_{n+1} + o(t). \end{aligned} \tag{2.34}$$

Thus $\mu = r'_{n+1}$ and

$$\mathscr{L}^{\xi}(T_n) = \mathscr{E}\left(\left(\sum_{k=n+1}^{\infty} r_k \right)^{-1} \right) \quad \text{if } \xi \in \Xi_n.$$

So (2.11) is true since by symmetry and by (2.30),

$$\begin{aligned} &\left| N^n p_n(tN^n, 0, \xi) - 1 \right| \\ &\quad \leq N^n \sum_{l=1}^{n} \mathbf{P}^0(T_l \geq tN^n) N^{-l} \\ &\quad \leq \sum_{l=0}^{n} N^{n-l} \exp\left(-tN^n \sum_{k=l+1}^{\infty} r_k \right) \to 0 \quad \text{as } n \to \infty. \end{aligned} \tag{2.35}$$

Also (2.12) holds by (2.30) and (2.26). □

2.4.3. *Case $X(t)$ critical.* Recall that a recurrent random walk on Ξ is called critical if

$$\log k \left[\log(N^k r_k) - \log(N^{k+1} r_{k+1}) \right] \text{ is bounded.} \tag{2.36}$$

This implies

$$\exists\, \varepsilon > 0\colon\ \varepsilon^{-1} > \frac{N^k r_k}{N^l r_l} > \varepsilon \qquad \forall\, l\ \forall\, k \in\,]l - \log l, l + \log l[. \tag{2.37}$$

LEMMA 2.6 (Critical random walk on Ξ). *A critical random walk on Ξ is diffusive in the sense of Definition* 2.1 *and* (c_n) *can be chosen as*

$$c_n = 2 \frac{N-1}{N+1} \frac{\log n}{N^n r_n}. \tag{2.38}$$

PROOF. Because of (2.37),

$$\sum_{k=n}^{\infty} r_k = \sum_{k=n}^{\infty} (N^k r_k) N^{-k} \sim \frac{N}{N-1} r_n. \tag{2.39}$$

Thus (a_n) can be chosen as [recall (2.5) and (2.6)]

$$a_n = \frac{(N-1)^2}{N+1} \sum_{k=1}^{n} \frac{1}{N^k r_k}, \tag{2.40}$$

since by (2.25)

$$\begin{aligned} A(n) &= (N-1) \sum_{k=1}^{n} \frac{N^{-k}}{\sum_{j=k}^{\infty} r_j + r_k/(N-1)} + \frac{N^{-n}}{\sum_{j=n}^{\infty} r_j + r_n/(N-1)} \\ &\sim (N-1) \sum_{k=1}^{n} \frac{N^{-k}}{r_k(N+1)/(N-1)} + \frac{N^{-n}}{r_n(N+1)/(N-1)} \qquad (2.41) \\ &= \frac{(N-1)^2}{N+1} \sum_{k=1}^{n} \frac{1}{N^k r_k} + \frac{N-1}{N+1} \frac{1}{N^n r_n} \sim a_n. \end{aligned}$$

Note that in particular,

$$\frac{A(n+1)}{A(n)} \to 1 \quad \text{as } n \to \infty. \tag{2.42}$$

Obviously $(c_n) \ll (a_n)$. By (2.25),

$$\begin{aligned} &\frac{1}{a_n} \sum_{m=[c_n N^n]}^{\infty} \left[p^{(m)}(0) - p^{(m)}(n)\right] \\ &\qquad = \frac{N-1}{a_n} \sum_{k=1}^{n} N^{-k} \frac{f_k^{[c_n N^n]}}{1-f_k} - \frac{1}{a_n} N^{-n} \frac{f_n^{[c_n N^n]}}{1-f_n}. \end{aligned} \tag{2.43}$$

Since

$$1 - f_k \sim \frac{N+1}{N-1} r_k, \tag{2.44}$$

we have

$$\sum_{k=1}^{\infty} \frac{1}{a_k} \frac{N^{-k}}{1-f_k} f_k^{c_k N^k} < \infty.$$

Since by recurrence $a_n \uparrow \infty$, applying Kronecker's lemma to (2.43) yields (2.14).

Now by (2.25),

$$\begin{aligned} N^n p^{([c_n N^n])}(0,0) &= (N-1) \sum_{k=1}^{\infty} N^{n-k} f_k^{[c_n N^n]} \\ &\sim (N-1) \sum_{k=1}^{\infty} N^{n-k} \exp\left(-2 \frac{N^k r_k}{N^n r_n} \log(n) N^{n-k}\right). \end{aligned} \tag{2.45}$$

Split up the sum into three parts

$$\sum_{k=1}^{n-\log n} + \sum_{k=n-\log n}^{n+\log n} + \sum_{k=n+\log n}^{\infty}$$

and observe that the summand attains a maximum of value less than or equal to $2/\varepsilon \log n$ at $k_0 = n + (\log(2\varepsilon) + \log\log n)/\log N$ and is monotone for $k < k_0$. Thus it is easily seen that (2.45) vanishes as $n \to \infty$, so (2.12) holds. Proving (2.13) is almost the same. First note by (2.25),

$$\begin{aligned}\frac{1}{a_n}\sum_{m=0}^{[c_nN^n]} p^{(m)}(0,0) &= \frac{N-1}{a_n}\sum_{k=1}^{\infty} N^{-k}\frac{1-f_k^{[c_nN^n]}}{1-f_k}\\ &\sim \frac{(N-1)^2}{a_n(N+1)}\sum_{k=1}^{\infty}\frac{1}{N^k r_k}\left[1-\exp\left(-2\log n\frac{r_k}{r_n}\right)\right].\end{aligned}$$

Now

$$\begin{aligned}\lim_{n\to\infty}&\frac{(N-1)^2}{a_n(N+1)}\sum_{k=1}^{n-\log n}\frac{1}{N^k r_k}\left[1-\exp\left(-2\log n\frac{r_k}{r_n}\right)\right]\\ &= \lim_{n\to\infty}\frac{(N-1)^2}{a_n(N+1)}\sum_{k=1}^{n-\log n}\frac{1}{N^k r_k} = 1,\end{aligned}$$

while $\sum_{n-\log n}^{n+\log n}$ and $\sum_{n+\log n}^{\infty}$ are shown to tend to 0 similarly as above. Finally (2.11) is obtained the same way as in the case $X(t)$ transient. □

2.4.4. *Key result on hitting times.* Up to now we have proved the following proposition:

PROPOSITION 2.7 (Diffusive random walks on Ξ). *Let $X(t)$ be a random walk on Ξ and let $X_n(t)$ be its restriction to Ξ_n.*

(a) *If $X(t)$ is transient and $s_n = GN^n$, then*

$$\mathbf{P}^{\xi_n}(\tau_n > ts_n) \to e^{-t} \quad \text{as } n\to\infty \tag{2.46}$$

uniformly in all sequences of starting points $(\xi_n)_{n\in\mathbb{N}}$ with $\xi_n \in \Xi_n$, $n\in\mathbb{N}$, such that $\|\xi_n\| \ge b_n$ for an arbitrary fixed sequence $b_n \uparrow \infty$.

(b) *If $X(t)$ is critical, $s_n = a_nN^n$, $\alpha\in[0,1]$ fixed and $(\xi_n)_{n\in\mathbb{N}}$ is a sequence with $\xi_n\in\Xi_n$, $n\in\mathbb{N}$, such that $(A(0,\xi_n))/(A(n)) \to \alpha$ as $n\to\infty$, then*

$$\mathbf{P}^{\xi_n}\left[\frac{\tau}{s_n} > t\right] \to \alpha \quad \text{as } n\to\infty, \tag{2.47}$$

$$\mathbf{P}^{\xi_n}\left[\frac{\tau_n}{s_n} > t\right] \to \alpha e^{-t} \quad \text{as } n\to\infty. \tag{2.48}$$

Recall $n(t)$ from (1.9).

COROLLARY 2.8 (Continuous time). *In the critical case the following continuous time version of* (2.47) *holds*:

$$\mathbf{P}^{\xi_{n(t)}}[\tau > t] \to \alpha \quad \text{as } t\to\infty.$$

PROOF. By (2.42) $s_n/(s_{n+1})$ is bounded and bounded away from 0 for n large enough. Thus (2.47) yields the assertion. □

3. Coalescing random walks. We introduce the notions of delayed coalescing random walks and instantaneously coalescing random walks and then give asymptotics for the number of surviving particles when scaling space and time properly. The main results are Propositions 3.2 and 3.4.

3.1. *Preparations.* Start with a system $\overline{X}(t) = (X(i,t))_{i=1,\dots,m}$ of independent copies of a random walk $X(t)$ on Ξ [resp. $X_n(i,t)$ on Ξ_n] starting at some initial points $\xi(i)$. Now think of $\overline{X}(t)$ as m particles moving on Ξ and let any two particles coalesce if they meet each other. That is, one of the two particles dies and the other goes on moving. Call this new process $\tilde{\eta}(t)$ the *system of instantaneously coalescing random walks*. Finally change the coalescence mechanism by not letting coalescence occur instantaneously, but at a constant rate $b > 0$. That is, a pair of particles coalesces after the particles have spent together an exponential waiting time with mean $1/b$. Call this new process a *system of delayed coalescing random walks* (with delay $1/b$) and denote it by $\eta(t)$. We are interested in $\eta(t)$ because of the mentioned duality relation. Since $\tilde{\eta}(t)$ is easier to handle, we first investigate this and then compare $\tilde{\eta}(t)$ with $\eta(t)$. By $\overline{X}_n$, $\eta_n(t)$, $\tilde{\eta}_n(t)$ and so forth we denote the corresponding objects on Ξ_n.

By forgetting the ordering of the particles we can regard $\eta(t)$ as a process on

$$\Phi := \left\{\varphi = (\varphi_\xi) \in \mathbb{N}_0^{\Xi}\colon\ \#\varphi := \sum_{\xi} \varphi_\xi < \infty\right\}, \tag{3.1}$$

where $\eta_\xi(t)$ is the number of particles at site ξ. The set Φ inherits the Tychonov topology from $(\mathbb{N}_0)^{\Xi}$. Note that $\eta(t)$ preserves $\Phi_m := \{\varphi \in \Phi\colon\ \#\varphi \le m\}$. For $\eta(0) \in \Phi_m$, $\eta(t)$ is the Markov process on Φ_m with generator $\mathscr{G}_m$ defined for $f \in C_b(\Phi_m)$ by

$$\begin{aligned}\mathscr{G}_m f(\varphi) = {}& \sum_{\xi,\zeta\in\Xi} \varphi_\xi \cdot a(\xi,\zeta)[f(\varphi - \mathbb{1}_\xi + \mathbb{1}_\zeta) - f(\varphi)] \\ & + \sum_{\xi\in\Xi} b\binom{\varphi_\xi}{2}[f(\varphi - \mathbb{1}_\xi) - f(\varphi)].\end{aligned} \tag{3.2}$$

[We use the convention $\binom{n}{k} = 0$ for $n < k$.]

On the other hand, $\tilde{\eta}(t)$ runs on

$$\tilde{\Phi} := \{\varphi \in \Phi\colon\ \varphi_\xi \in \{0,1\}\,\forall\,\xi\}. \tag{3.3}$$

The process $\tilde{\eta}(t)$ preserves $\tilde{\Phi}_m := \tilde{\Phi} \cap \Phi_m$ and on this has generator $\mathscr{H}_m$ defined for $f \in C_b(\tilde{\Phi}_m)$ by

$$\mathscr{H}_m f(\varphi) = \sum_{\xi,\zeta\in\Xi} \varphi_\xi \cdot a(\xi,\zeta)[f((\varphi - \mathbb{1}_\xi + \mathbb{1}_\zeta)\wedge 1) - f(\varphi)]. \tag{3.4}$$

3.2. *Scaling properties of* $\widetilde{\eta}(t)$ *on* Ξ. We first look into the case $X(t)$ critical and then consider the case $X(t)$ transient. Hence assume now $a(\cdot,\cdot)$ to be critical.

We fix $m \in \mathbb{N}$ and start $\widetilde{\eta}(t)$ with particles at sites $\xi_{n,1}, \dots, \xi_{n,m}$, that is, in

$$\varphi_n := \mathbb{1}_{\xi_{n,1}} + \cdots + \mathbb{1}_{\xi_{n,m}}, \tag{3.5}$$

such that [recall $A(n)$ from (2.6) and (2.41)]

$$\frac{A(\xi_{n,i}, \xi_{n,j})}{A(n)} \to \alpha \quad \text{as } n \to \infty \ \forall\, i \neq j. \tag{3.6}$$

In order to formulate the main result of this subsection we shall need the following definition.

DEFINITION 3.1 (Pure death process). With $(D_t)_{t\geq 0}$ we denote the *nonlinear pure death process* on $\mathbb{N}$ that jumps from m to $m-1$ at rate $\binom{m}{2}$. By

$$q_t(m;k) = \mathbf{P}^m(D_t = k) \tag{3.7}$$

we denote its transition probabilities.

Note that $q_t(m;m) = \exp[-\binom{m}{2}t]$ and recall $\hat{\alpha} = -\log\alpha$.

PROPOSITION 3.2 (Scaling limit, infinite case).

$$\mathbf{P}^{\varphi_n}\left[\#\widetilde{\eta}(s_n) = k\right] \to q_{\hat{\alpha}}(m;k) \quad \textit{as } n \to \infty. \tag{3.8}$$

We introduce the notation

$$\begin{aligned} \tau(i,j) &:= \inf\{t \geq 0\colon\ X(i,t) = X(j,t)\},\\ \overline{\tau} &:= \min_{i\neq j} \tau(i,j). \end{aligned} \tag{3.9}$$

In view of Corollary 2.8 it suffices to let $t \to \infty$ along the fixed sequence $t_n = s_n$. Our main goal for proving Proposition 3.2 is then to establish that the $\binom{m}{2}$ pairs of particles happen to coalesce asymptotically independently in the infinite case and that the "meeting probability" is given by our quantity α. Namely, we show the following lemma:

LEMMA 3.3.

$$\mathbf{P}^{\varphi_n}\left[\overline{\tau} \leq s_n\right] \to 1 - \alpha^{\binom{m}{2}} \quad \textit{as } n \to \infty. \tag{3.10}$$

Following the lines of the proof of Theorem 5 of Cox and Griffeath (1986), an induction argument then proves the proposition. We will not repeat the latter argument here.

PROOF OF LEMMA 3.3. We first rewrite the relation (3.10) in a more tractable form using Proposition 2.7. Namely [recall a_n and r_n from (1.5) and (2.6)],

$$\mathbf{P}^{\varphi_n}\left[\overline{\tau} \le \frac{1}{r_n}\right] \to 1 - \alpha^{\binom{m}{2}} \quad \text{as } n \to \infty. \tag{3.11}$$

To see this equivalence we argue as follows. Note that s_n in Proposition 2.7 can be replaced by $1/r_n$ since we can choose (n'): $s_{n'} \le 1/r_n$ and $n - n' = o(\log n)$, so (2.36) implies $a_{n'}/a_n \to 1$ as $n \to \infty$. Thus for $\gamma \in [\alpha, 1]$, Proposition 2.7 asserts

$$\mathbf{P}^{\xi_n}[\tau > t(\gamma, n)] \to \frac{\alpha}{\gamma} \quad \text{as } n \to \infty, \tag{3.12}$$

where we put $t(\gamma, n) := 1/r_{f_\gamma(n)}$. w.l.o.g. we assume $f_1(n) = n$.

So we concentrate on showing (3.11). Again by (2.36) for any $\gamma \in [\alpha, 1]$ there exist sequences $d(\gamma, n)$, $e(\gamma, n)$ such that

$$\lim_{n\to\infty}[f_\gamma(n) - d(\gamma, n)] = \lim_{n\to\infty}[e(\gamma, n) - f_\gamma(n)] = \infty$$

and

$$\lim_{n\to\infty} \frac{A(d(\gamma, n))}{A(n)} = \lim_{n\to\infty} \frac{A(e(\gamma, n))}{A(n)} = \gamma.$$

These can be assumed to be increasing in γ.

Let

$$\Xi(\gamma, n) := \{\xi \in \Xi\colon \ \|\xi\| \in [d(\gamma, n), e(\gamma, n)]\}.$$

Note that Proposition 2.7 is valid uniformly in all sequences $(\xi_n)_{n\in\mathbb{N}}$ with $\xi_n \in \Xi(\alpha, n)$.

Now

$$\mathbf{P}^{\xi_n}\left[X(t(\gamma, n)) \in \Xi(\gamma, n)\right] \to 1 \quad \text{as } n \to \infty \tag{3.13}$$

since

$$\begin{aligned} &\mathbf{P}^{\xi_n}\left[\|X(t(\gamma, n))\| \ge e(\gamma, n)\right] \\ &\quad \le \mathbf{P}^{\xi_n}\left[T_{e(\gamma,n)} \le t(\gamma, n)\right] \\ &\quad = \exp\left(-t(\gamma, n) \sum_{k\ge e_\gamma(n)} r_k\right) \to 0 \quad \text{as } n \to \infty \end{aligned} \tag{3.14}$$

by (2.39). The opposite direction works similarly.

Denote by $\varepsilon(n)$ any quantity tending to 0 as $n \to \infty$. We shall make use of the following auxiliary equations:

$$\begin{aligned} \int_\alpha^1 \mathbf{P}^{\varphi_n}\big[\overline{\tau} = \tau(1, 2) \in dt(n, \gamma),\ & X(2, t(n, \gamma)) \\ & - X(3, t(n, \gamma)) \notin \Xi(n, \gamma)\big] = \varepsilon(n) \end{aligned} \tag{3.15}$$

and

$$\int_\alpha^1 \mathbf{P}^{\varphi_n}\big[\overline{\tau}=\tau(1,2)\in dt(n,\gamma),\ X(4,t(n,\gamma)) - X(3,t(n,\gamma))\notin \Xi(n,\gamma)\big]=\varepsilon(n). \tag{3.16}$$

We prove only (3.15), since the proof of (3.16) is even simpler:

$$\begin{aligned}
&\int_\alpha^1 \mathbf{P}^{\varphi_n}\big[\overline{\tau}=\tau(1,2)\in dt(n,\gamma),\ X(2,t(n,\gamma))-X(3,t(n,\gamma))\notin \Xi(n,\gamma)\big]\\
&\quad=\int_\alpha^1 \sum_{\xi\in\Xi}\mathbf{P}^{\varphi_n}\big[\overline{\tau}=\tau(1,2)\in dt(n,\gamma),\ X(2,t(n,\gamma))=\xi\big]\\
&\qquad\qquad\times \mathbf{P}\big[X(3,t(n,\gamma))-\xi\notin \Xi(n,\gamma)\big]\quad \text{[by symmetry and (3.13)]}\\
&\quad=\int_\alpha^1 \sum_{\xi\in\Xi}\mathbf{P}^{\varphi_n}\big[\overline{\tau}=\tau(1,2)\in dt(n,\gamma),\ X(2,t(n,\gamma))=\xi\big]\\
&\qquad\qquad\times \mathbf{P}\big[X(3,t(n,\gamma))\notin \Xi(n,\gamma)\big]+\varepsilon(n)\\
&\quad=\varepsilon(n)\quad \text{(by dominated convergence).}
\end{aligned}$$

Now we put the pieces together:

$$\begin{aligned}
&\mathbf{P}^{\varphi_n}\bigg[\tau(i,j)\le \frac{1}{r_n}\bigg]\\
&\quad=\mathbf{P}^{\varphi_n}\bigg[\overline{\tau}=\tau(i,j)\le\frac{1}{r_n}\bigg]\\
&\qquad+\sum_{\{k,l\}\ne\{i,j\}}\int_{t(\alpha,n)}^{t(1,n)}\mathbf{P}^{\varphi_n}\big[\overline{\tau}=\tau(k,l)\in dt,\ \tau(i,j)\le t(1,n)\big]+\varepsilon(n).
\end{aligned}\tag{3.17}$$

We substitute to change the domain of integration to $[\alpha,1]$. We then condition the integrand on $(X(i,t),X(j,t))$ and apply the Markov property. With (3.13) and (3.12) we get that the integral term in (3.17) equals

$$\begin{aligned}
&\int_\alpha^1 \mathbf{P}^{\varphi_n}\big[\overline{\tau}=\tau(k,l)\in dt(\gamma,n),\\
&\qquad X(i,t(\gamma,n))-X(j,t(\gamma,n))\in\Xi(\gamma,n),\ \tau(i,j)\le t(1,n)\big]+\varepsilon(n).
\end{aligned}\tag{3.18}$$

Apply (3.15) and (3.16) to see that this in turn equals

$$\begin{aligned}
&\int_\alpha^1 \sum_{\xi-\zeta\in\Xi(\gamma,n)}\mathbf{P}^{\varphi_n}\big[\overline{\tau}=\tau(k,l)\in dt(\gamma,n),\ X(i,t(\gamma,n))=\xi,\ X(j,t(\gamma,n))=\zeta\big]\\
&\qquad\qquad\times\mathbf{P}^{(\xi,\zeta)}\big[\tau(1,2)\le t(1,n)-t(\gamma,n)\big]+\varepsilon(n)\\
&\quad=\int_\alpha^1 \mathbf{P}^{\varphi_n}\big[\overline{\tau}=\tau(k,l)\in dt(\gamma,n)\big](1-\gamma)+\varepsilon(n).
\end{aligned}$$

Integration by parts and summation over all pairs $\{i, j\}$ in (3.17) yields

$$\begin{aligned}\binom{m}{2}(1-\alpha) &= \mathbf{P}^{\varphi_n}\left[\bar{\tau} \le \frac{1}{r_n}\right] \\ &\quad + \left(\binom{m}{2} - 1\right)\int_\alpha^1 \mathbf{P}^{\varphi_n}\left[\bar{\tau} \le t(\gamma, n)\right] d\gamma + \varepsilon(n).\end{aligned} \tag{3.19}$$

A contraction argument [compare again Cox and Griffeath (1986)] now shows

$$\mathbf{P}^{\varphi_n}\left[\bar{\tau} \le \frac{1}{r_n}\right] \to 1 - q_{\hat{\alpha}}(m; m) = 1 - \alpha^{\binom{m}{2}t} \quad \text{as } n \to \infty.$$

So we are done. □

3.3. *Scaling properties of* $\widetilde{\eta}_n(t)$ *on* Ξ_n. We now turn to finite systems. Here also particles coalesce asymptotically independently but the "hitting probabilities" are different.

PROPOSITION 3.4 (Scaling limit, finite case).

$$\mathbf{P}^{\varphi_n}\left[\#\widetilde{\eta}_n(ts_n) = k\right] \to q_{2t+\hat{\alpha}}(m; k) \quad \textit{as } n \to \infty. \tag{3.20}$$

PROOF. We prove the statement for $\alpha = 1$. The general case then can be obtained from this as follows. As in the proof of Lemma 3.3 we can choose a sequence (s'_n) with $s'_n/s_n \to 0$ as $n \to \infty$ slowly enough that

$$\begin{aligned}&\mathbf{P}^{\varphi_n}\big[X(i, u) \in \Xi_n \ \forall\, u \le s'_n, \ \forall\, i \\ &\qquad \text{and } X(i, s'_n) - X(j, s'_n) \in \Xi(1, n), \ \forall\, i \ne j\big] \to 1 \quad \text{as } n \to \infty\end{aligned} \tag{3.21}$$

and

$$\mathbf{P}^{\varphi_n}\left[\#\widetilde{\eta}(ts_n) = k\right] \to q_{\hat{\alpha}}(m; k) \quad \text{as } n \to \infty.$$

Since given the event in (3.21), $\widetilde{\eta}_n(s'_n) = \widetilde{\eta}(s'_n)$, and since $ts_n \sim ts_n - s'_n$, we have

$$\begin{aligned}\mathbf{P}^{\varphi_n}\left[\#\widetilde{\eta}_n(ts_n) = k\right] &= \sum_{l=k}^{m} \mathbf{P}^{\varphi_n}\left[\#\widetilde{\eta}(ts_n) = l\right] q_{2t}(l; k) + \varepsilon(n) \\ &= \sum_{l=k}^{m} q_{\hat{\alpha}}(m; l) q_{2t}(l; k) + \varepsilon(n) = q_{2t+\hat{\alpha}}(m; k) + \varepsilon(n).\end{aligned} \tag{3.22}$$

The last equality is of course the Chapman–Kolmogorov equality.

Hence we assume now $\alpha = 1$. Note that $X(i, t) - X(j, t)$ is a random walk running at double speed. So by Proposition 2.7 the analogue of (3.12) is [recall $\tau_n(i, j)$ and $\bar{\tau}_n$ are the finite objects corresponding to those defined in (3.9)]

$$\mathbf{P}^{\varphi_n}\left[\tau_n(i, j) \le ts_n\right] \to 1 - e^{-2t} \quad \text{as } n \to \infty. \tag{3.23}$$

Thus we replace α by e^{-2t} in the proof of Lemma 3.3 to obtain

$$\mathbf{P}^{\varphi_n}\left[\overline{\tau}_n \le ts_n\right] \to 1 - \exp\left(-2t\binom{m}{2}\right) \quad \text{as } n \to \infty. \tag{3.24}$$

Now the induction argument cited above yields

$$\mathbf{P}^{\varphi_n}\left[\#\widetilde{\eta}_n(ts_n) = k\right] \to q_{2t}(m;k) \quad \text{as } n \to \infty. \tag{3.25}$$

□

3.4. *Case a recurrent, comparison of $\eta(t)$ and $\widetilde{\eta}(t)$.* Let $X(t)$ (or a) be recurrent. We show that in our space and time scaling delayed and instantaneously coalescing random walks η and $\widetilde{\eta}$ (resp. η_n and $\widetilde{\eta}_n$) are equivalent in the following sense:

For $\varphi \in \Phi$, let $\varphi^* = \varphi \wedge 1$ denote the projection to $\widetilde{\Phi}$ and $\eta^{\varphi}(t)$ [resp. $\widetilde{\eta}^{\varphi^*}(t)$], the systems started in φ (resp. φ^*). Fix m and m^* and choose (φ_n) such that $\#\varphi_n = m$, $\#\varphi_n^* = m^*$ and (φ_n^*) is an α-spaced sequence in the sense of (3.6).

LEMMA 3.5 (Comparison). *Under these conditions*,

$$\mathbf{P}\left[\widetilde{\eta}^{\varphi_n^*}(s_n) = \eta^{\varphi_n}(s_n)\right] \to 1 \quad \textit{as } n \to \infty, \tag{3.26}$$

$$\mathbf{P}\left[\widetilde{\eta}_n^{\varphi_n^*}(ts_n) = \eta_n^{\varphi_n}(ts_n)\right] \to 1 \quad \textit{as } n \to \infty,\ t > 0. \tag{3.27}$$

PROOF. We shall only show (3.26) since (3.27) is similar. Let

$$T_i^n = \inf\left\{t \ge 0\colon \#\widetilde{\eta}^{\varphi_n^*}(t) = m^* - i\right\}, \qquad i = 0, 1, \dots, m^* - 1,$$

be the time points of coalescence and note that

$$T_{i+1}^n - T_i^n \to \infty \quad \text{as } n \to \infty,\ \mathbf{P}\text{-a.s.}$$

Hence by recurrence the particles that meet at time T_i^n coalesce in η^{φ_n} until time T_{i+1}^n asymptotically, $\mathbf{P}$-a.s. □

Combining Propositions 3.2 and 3.4 and Lemma 3.5, we have proved the following proposition.

PROPOSITION 3.6 (Scaling limits).

$$\mathscr{L}^{\varphi_{n(t)}}(\#\eta(t)) \Rightarrow \mathscr{L}^{m^*}(D_{\hat{\alpha}}) \quad \textit{as } t \to \infty,$$
$$\mathscr{L}^{\varphi_n}(\#\eta_n(ts_n)) \Rightarrow \mathscr{L}^{m^*}(D_{2t+\hat{\alpha}}) \quad \textit{as } n \to \infty,\ \textit{for } t > 0 \textit{ fixed}.$$

3.5. *Case a transient, comparison of $\eta(t)$ and $\eta_n(t)$.* We now look into the case a transient. The comparison lemma does not hold here because it depends heavily on the recurrence property of a. We used that once a pair meets, it meets infinitely often in the large time scale and finally coalesces. So we have to do some more subtle computations now in the transient case.

Fix a sequence $t_n \uparrow \infty$, $t_n \ll (\sum_{k>n} r_k)^{-1}$. Then [by (2.30)] $t_n \ll N^{-n}$ and

$$\mathbf{P}[X_n(t) = X(t),\ \forall\ t \le t_n] \to 1 \quad \text{as } n \to \infty. \tag{3.28}$$

Let $\tau_n^{(0)} = 0$ and

$$\tau_n^{(i+1)} = \inf\{t > \tau_n^{(i)} + t_n:\ X_n(t) = 0\}. \tag{3.29}$$

Since

$$\sup_{x \in \Xi_n} \mathbf{E}^x[G(0, X_{t_n})] = \mathbf{E}^0[G(0, X_{t_n})] \to 0 \quad \text{as } n \to \infty, \tag{3.30}$$

by Proposition 2.7 we get

$$\mathscr{L}\left[\frac{\tau_n^{(i+1)} - \tau_n^{(i)}}{G|\Xi_n|}\right] \Rightarrow \mathscr{E}(1) \quad \text{as } n \to \infty. \tag{3.31}$$

Let $B(t)$ a Poisson process with rate 1. Then for $t > 0$,

$$\mathscr{L}\left[\max\left\{k\text{: } \frac{\tau_n^{(k)}}{G|\Xi_n|} \le t\right\}\right] \Rightarrow \mathscr{L}[B(t)] \quad \text{as } n \to \infty. \tag{3.32}$$

Recall

$$V = \mathbf{E}^0\left[\exp\left(-\frac{1}{2}\int_0^\infty \mathbb{1}_{\{X_s = 0\}}\, ds\right)\right] \tag{3.33}$$

and let

$$p_\varphi(k) := \lim_{t \to \infty} \mathbf{P}^\varphi[\#\eta(t) = k]. \tag{3.34}$$

Note that

$$V = p_{(0,0)}(2). \tag{3.35}$$

By (3.28) and (3.32) we get

$$\lim_{n \to \infty} \mathbf{P}^{(\zeta,\xi)}[\#\eta_n(tGN^n) = 1] = 1 - p_{(\zeta,\xi)}(2)\exp(-2t(1 - V)). \tag{3.36}$$

Now proceeding as above we get that the pairs of particles coalesce (asymptotically) independently. Thus if we put

$$s_n = \frac{G}{1 - V} N^n, \tag{3.37}$$

we obtain the following proposition.

PROPOSITION 3.7.

$$\mathbf{P}^\varphi[\#\eta_n(ts_n) = k] \to \sum_l p_\varphi(l) q_{2t}(l;k) \quad \textit{as } n \to \infty. \tag{3.38}$$

4. Proofs of Theorems 1, 3 and 4.

4.1. *Proof of Theorems* 1 *and* 4. Since parts (a) are immediate consequences of parts (b), we will only show (b). We first look into the special case where we start at the product measure π_θ and where $g(x) = bx(1 - x)$, $b > 0$.

4.1.1. *Special case $g(x) = bx(1-x)$ and product measure.* Since we will have to work with various diffusion coefficients g, we add g or b as a superscript where necessary. Let now $\eta(t)$ be a system of coalescing random walks with delay $1/b$ and let

$$z^{\varphi} := \prod_{\xi\in\Xi} (z_{\xi})^{\varphi_{\xi}}, \qquad z \in [0,1]^{\Xi},\ \varphi \in \Phi.$$

Our main tool is the following duality relation between mixed moments of interacting diffusions and delayed coalescing random walks:

$$\mathbf{E}^{z}\big[(\mathbb{X}^{b}(t))^{\varphi}\big] = \mathbf{E}^{\varphi}\big[z^{\eta(t)}\big], \tag{4.1}$$

which is also true for finite systems. For a proof, see Shiga [(1980), Lemma 2.3].

Since the state space is compact, it suffices to show convergence of (mixed) moments. Thus we fix $\varphi = k_1 \mathbb{1}_{\xi_1} + \cdots + k_{m^*} \mathbb{1}_{\xi_{m^*}} \in \Phi$, $m^* \in \mathbb{N}$, $k_1, \dots, k_{m^*} \in \mathbb{N}$, $\xi_i \neq \xi_j$, that is, a point in Φ with k_j particles at site ξ_j. Let $\varphi_n = S^{-1}_{f_\alpha(n)}\varphi$ be the spaced version of φ. We have to show

$$\begin{aligned} \mathbf{E}^{\pi_\theta}\big[\big({}^{f_\alpha}\mathbb{X}^{b}(t)\big)^{\varphi}\big] &= \mathbf{E}^{\pi_\theta}\big[\big(\mathbb{X}^{b}(t)\big)^{\varphi_{n(t)}}\big] \\ &\to \mathbf{E}^{\theta}\big[(Y_{\hat\alpha})^{m^*}\big] \quad \text{as } t \to \infty, \end{aligned} \tag{4.2}$$

$$\begin{aligned} \mathbf{E}^{\pi_\theta}\big[\big({}^{f_\alpha}\mathbb{X}^{b}_{n}(n)\big)^{\varphi}\big] &= \mathbf{E}^{\pi_\theta}\big[\big(\mathbb{X}^{b}_{n}(ts_n)\big)^{\varphi_n}\big] \\ &\to \mathbf{E}^{\theta}\big[(Y_{2t+\hat\alpha})^{m^*}\big] \quad \text{as } n \to \infty. \end{aligned} \tag{4.3}$$

By (4.1) and Proposition 3.6 the l.h.s. of (4.2) equals

$$\begin{aligned} \int \mathbf{E}^{\varphi_{n(t)}}\big[z^{\eta(t)}\big]\pi_\theta(dz) &= \mathbf{E}^{\varphi_{n(t)}}\big[\theta^{\#\eta(t)}\big] \\ &\to \mathbf{E}^{m^*}\big[\theta^{D_{\hat\alpha}}\big] = \mathbf{E}^{\theta}\big[(Y_{\hat\alpha})^{m^*}\big] \quad \text{as } t \to \infty. \end{aligned}$$

The last equality is a well known duality between the Fisher–Wright diffusion and the pure death process introduced in Definition 3.1. The proof of (4.3) is fairly similar.

4.1.2. *Generalization to ergodic measures.* Here we want to generalize the result to ergodic start measures μ with intensity θ. We do so by coupling techniques; that is, we show that two versions $\mathbb{X}^1$ and $\mathbb{X}^2$ of our interacting system with ergodic initial laws μ and ν with the same intensity θ can be defined on one probability space such that $\mathbf{E}[|x^1_0(t) - x^2_0(t)|] \to 0$ as $t \to \infty$. Define the four-valued process $(\mathbb{X}^1, \mathbb{X}^2, \mathbb{X}^1_n, \mathbb{X}^2_n)$ as the solution of

$$\begin{aligned} dx^{i}_{\xi}(t) &= \sum_{z\in\Xi} a(\xi,\zeta)(x^{i}_{\zeta}(t) - x^{i}_{\xi}(t))\,dt \\ &\quad + \sqrt{bx^{i}_{\xi}(1-x^{i}_{\xi})}\,dW_{\xi}(t), \qquad i = 1, 2, \end{aligned} \tag{4.4}$$

$$
\begin{aligned}
dx^i_{n,\xi}(t) &= \sum_{z\in\Xi_n} a(\xi,\zeta)(x^i_{n,\zeta}(t) - x^i_{n,\xi}(t))\,dt \\
&\quad + \sqrt{bx^i_{n,\xi}(1-x^i_{n,\xi})}\,dW_\xi(t), \qquad i = 1,2,
\end{aligned} \tag{4.5}
$$

with *one* set of Brownian motions and the initial common law given by

$$\mathscr{L}(\mathbb{X}^1(0), \mathbb{X}^2(0)) = \mu \otimes \nu$$

and

$$\left(\mathbb{X}^1_n(0), \mathbb{X}^2_n(0)\right) = \left(\mathbb{X}^1(0), \mathbb{X}^2(0)\right)\big|_{\Xi_n} \qquad (\mu\otimes\nu)\text{-a.e.}$$

Here μ and ν are spatially ergodic with same intensity θ. Let $\Delta_\xi(t) = x^1_\xi(t) - x^2_\xi(t)$, $\Delta_{n,\xi}(t) = x^1_{n,\xi}(t) - x^2_{n,\xi}$ and $\Delta^n_\xi(t) = x^1_{n,\xi}(t) - x^1_\xi(t)$.

We will rely on the following lemma, which is due to Cox and Greven [(1994a), Lemma 4] in the case a transient and due to Fleischmann and Greven [(1994), Proposition 5.11] in the case a recurrent. (Fleischmann and Greven only deal with the case a critical, but the proof they give actually works for any a recurrent. In fact, a slight modification of their proof yields a unified approach to both cases, a recurrent and a transient.)

LEMMA 4.1 (Successful coupling, infinite systems). *Assume $a(\cdot,\cdot)$ to be either transient or recurrent. Then*

$$\mathbf{E}\left[|\Delta_0(t)|\right] \to 0 \quad \text{as } t\to\infty. \tag{4.6}$$

This yields the analogue of (4.2) if we put $\nu = \pi_\theta$. So we are done with the infinite case.

We polish off the finite case by deriving (based on Lemma 4.1) the following lemma.

LEMMA 4.2 (Successful coupling, finite systems). *Under the same conditions as in Lemma* 4.1,

$$\mathbf{E}\left[|\Delta_{n,0}(ts_n)|\right] \to 0 \quad \text{as } n\to\infty. \tag{4.7}$$

PROOF. Since the infinite systems can be coupled successfully, we have to show that the finite and the infinite systems do not diverge for sufficiently large time and that finite systems stay close once they get close. Fix a sequence $t_m \uparrow \infty$ such that $t_m \ll \left(\sum_{k>m} r_k\right)^{-1}$ [recall r_k from (1.5)]. Then

$$\sup_{n\ge m} \mathbf{E}\left[|\Delta^n_\xi(t_m)|\right] \to 0 \quad \text{as } m\to\infty. \tag{4.8}$$

To see this we may proceed as Yamada and Watanabe (1971). We approximate the $|\cdot|$ function by functions $f_n(x) = \sqrt{1/n + x^2}$ to which the Itô formula can be applied and obtain

$$d|\Delta^n_\xi(t)| = \operatorname{sgn}(\Delta^n_\xi(t))\,d\Delta^n_\xi(t). \tag{4.9}$$

Then

$$\begin{aligned} d\,\mathbf{E}\,|\Delta_\xi^n(t)| &\le \mathbf{E}\biggl[\sum_{\zeta\in\Xi_n} a(\xi,\zeta)\bigl(|\Delta_\zeta^n(t)| - |\Delta_\xi^n(t)|\bigr)\biggr]\,dt \\ &\quad + \mathbf{E}\biggl[\sum_{\zeta\notin\Xi_n} a(\xi,\zeta)\bigl(|x_\zeta(t)| + |x_{\xi_{|\Xi_n}}(t)|\bigr)\biggr]\,dt. \end{aligned} \tag{4.10}$$

The first term vanishes by translation invariance (Ξ_n is a subgroup of Ξ) and the second term is bounded by $(2\sum_{k>n} r_k)\,dt$.

By Lemma 4.1 the infinite systems are close at time t_n, that is, $\mathbf{E}\,\bigl[|\Delta_0(t_n)|\bigr] = \varepsilon(n)$, and so are the finite systems. Hence it is enough to show

$$d\,\mathbf{E}\,\bigl[|\Delta_{n,\,\xi}(t)|\bigr] \le 0. \tag{4.11}$$

This is, however, true since, as above,

$$d\,\mathbf{E}\,\bigl[|\Delta_{n,\,\xi}(t)|\bigr] \le \mathbf{E}\biggl[\sum_{\zeta\in\Xi_n} a(\xi,\zeta)\bigl(|\Delta_{n,\,\zeta}(t)| - |\Delta_{n,\,\xi}(t)|\bigr)\biggr]\,dt = 0. \tag{4.12}$$

□

4.1.3. *Generalization to admissible* $g(x)$. Finally we generalize the diffusion coefficient. Fix an admissible g [recall (1.2)]. The idea is to sandwich g between two Fisher–Wright-type diffusion coefficients. We will then infer that the moments are also sandwiched by quantities that have the same limiting behavior according to the discussion in the last two subsections.

Fix $\frac{1}{2} > \varepsilon > 0$ and φ and let

$$\begin{aligned} f(x) &= x(1-x), \\ f^\varepsilon(x) &= [(x-\varepsilon)(1-x-\varepsilon)]^+. \end{aligned}$$

Choose $b,\ b^\varepsilon > 0$ such that

$$g^\varepsilon := b^\varepsilon f^\varepsilon \le g \le bf.$$

Denote by $\mathbb{X}^g(t)$, $\mathbb{X}^{g^\varepsilon}(t)$ and $\mathbb{X}^{bf}(t)$ the solutions of (1.1) driven by g, g^ε and bf, respectively, and with the same initial law μ. The crucial point is the comparison of the mixed moments of these,

$$\mathbf{E}^\mu\bigl[\bigl(\mathbb{X}^{g^\varepsilon}(t)\bigr)^\varphi\bigr] \le \mathbf{E}^\mu\bigl[\bigl(\mathbb{X}^{g}(t)\bigr)^\varphi\bigr] \le \mathbf{E}^\mu\bigl[\bigl(\mathbb{X}^{bf}(t)\bigr)^\varphi\bigr] \qquad \forall\, t\ge 0, \tag{4.13}$$

which is due to Cox, Fleischmann and Greven [(1996), Theorem 1].

We introduce the linear map

$$\begin{aligned} L^\varepsilon\colon\ [\varepsilon, 1-\varepsilon]^\Xi &\to [0,1]^\Xi, \\ (x_\xi) &\mapsto \left(\frac{x_\xi - \varepsilon}{1-2\varepsilon}\right) \end{aligned}$$

and its inverse H^ε. Let $\mu^\varepsilon := H^\varepsilon\mu$ and note that $\langle x_0, \mu\rangle - \langle x_0, \mu^\varepsilon\rangle = O(\varepsilon)$. Observe that the coupling of the last subsection [in particular, (4.11)] adapted to this setting yields

$$\mathbf{E}^{\mu^\varepsilon}\bigl[\bigl(\mathbb{X}^{g^\varepsilon}(t)\bigr)^\varphi\bigr] - \mathbf{E}^\mu\bigl[\bigl(\mathbb{X}^{g^\varepsilon}(t)\bigr)^\varphi\bigr] = O(\varepsilon). \tag{4.14}$$

Note that $L^{\varepsilon}\mathbb{X}^{g^{\varepsilon}}(t)$ is again of the Fisher–Wright-type for $\mathbb{X}(0)$ concentrated on $[\varepsilon, 1-\varepsilon]^{\Xi}$. Observe that $(H^{\varepsilon}(z))_0 - z_0 = O(\varepsilon)$, where the O constants only depend on $m = \#\varphi$. So the discussion of the last two subsections yields

$$\limsup_{\varepsilon\to 0}\limsup_{t\to\infty}\sup_{\#\varphi=m}\left|\mathbf{E}^{\mu}\left[\left(\mathbb{X}^{g^{\varepsilon}}(t)\right)^{\varphi}\right] - \mathbf{E}^{\mu}\left[\left(\mathbb{X}^{bf}(t)\right)^{\varphi}\right]\right| = 0. \tag{4.15}$$

This finishes the proof.

4.2. *Proof of Theorem* 3. Assume now $a(\cdot,\cdot)$ to be transient. Since the coupling of finite systems is successful (Lemma 4.2), we may assume

$$\mathscr{L}[\mathbb{X}(0)] = \pi_{\theta}.$$

Recall the definition of p_{φ} from (3.34) and note that

$$\mathbf{E}^{\nu_{\theta}}[z^{\varphi}] = \sum_k p_{\varphi}(k)\theta^k. \tag{4.16}$$

We proceed as above and use Proposition 3.7 to conclude Theorem 3:

$$\begin{aligned}\mathbf{E}^{\pi_{\theta}}\left[\left(\mathbb{X}_n(ts_n)\right)^{\varphi}\right] &= \mathbf{E}^{\varphi}\left[\theta^{\#\eta_n(ts_n)}\right]\\ &= \sum_l p_{\varphi}(l)\sum_k q_{2t}(l;k)\theta^k\\ &= \sum_l p_{\varphi}(l)\,\mathbf{E}^l\left[\theta^{D_{2t}}\right]\\ &= \sum_l p_{\varphi}(l)\int Q_{2t}(\theta, d\rho)\rho^l\\ &= \int Q_{2t}(\theta, d\rho)\,\mathbf{E}^{\nu_{\rho}}[z^{\varphi}].\end{aligned} \tag{4.17}$$

□

5. Proofs of Theorems 2 and 5. We only consider the case $g(x) = x(1-x)$ and $\mathscr{L}[\mathbb{X}(0)]$ product measure, since the generalizations work as in Section 4. Again we first have to do some random walk analysis. We start with the construction of the limit object of space- and time-scaled random walks on Ξ. From this we conclude parts (a) and (b) of Theorems 2 and 5. Then we obtain (c) via a duality to the discrete time nonlinear death process of Definition 3.1.

5.1. *Limit process of scaled random walks.* In this subsection we "extend Ξ (resp. Ξ_n) to the left," that is, by points of short distance, to Γ (resp. Γ') defined below. On these extended groups we will define the weak limits of rescaled random walks on Ξ (resp. Ξ_n).

DEFINITION 5.1. Let

$$\Gamma := \{\delta = (\delta_k)_{k\in\mathbb{Z}}\colon\ \delta_k \in \{0,\dots,N-1\},\ \|\delta\| < \infty\}, \tag{5.1}$$

$$\Gamma_{-n} := \{\delta\in\Gamma\colon\ \delta_k = 0\ \forall\, k \le -n\}, \tag{5.2}$$

where $\|\delta\| := \inf\{k \in \mathbb{Z}:\ \delta_k = 0\ \forall\, l > k\}$. Then Γ is an Abelian group with addition componentwise modulo N and Γ inherits the product topology from $\{0, \dots, N-1\}^{\mathbb{Z}}$.

The finite objects will be indicated by a prime and are defined as

$$\Gamma' := \{\delta \in \Gamma:\ \|\delta\| \le 0\},$$

$$\Gamma'_{-n} := \{\delta \in \Gamma_{-n}:\ \|\delta\| \le 0\}.$$

Further let μ (resp. μ') be the Haar measures on Γ (resp. Γ') normed to $\mu(\Gamma') = \mu'(\Gamma') = 1$ (sic), that is, the weak limits of N^{-n} times counting measure on Γ_{-n} (resp. Γ'_{-n}) as $n \to \infty$.

The shift operators S_k [recall (1.10)] naturally extend to these objects. Note that we may identify Ξ with Γ_0 and observe

$$S_n(\Xi) = \Gamma_{-n},$$

$$S_n(\Xi_n) = \Gamma'_{-n}.$$

Since most of what follows is the same for the finite and infinite objects, we suppress the prime where possible and only stress the occurring differences.

We obtain random walks $\gamma_n(t)$ on Γ by shifting a random walk $X(t)$ on Ξ and rescaling time

$$\gamma_n(t) := S_n(X(t(Nc)^{n+1})). \tag{5.3}$$

Intuitively we extend $X(t)$ "to the left" by allowing jumps of short distances at high rates.

In the same way we obtain the system of instantaneously coalescing random walks $\widetilde{\beta}_n$ on Γ:

$$\widetilde{\beta}_n(t) := S_n(\widetilde{\eta}(t(Nc)^{n+1})). \tag{5.4}$$

Denote by G_n the generator of γ_n defined on $C(\Gamma_{-n})$, the set of continuous functions on Γ_{-n}. We will identify $C(\Gamma_{-n})$ with $\widehat{C}(\Gamma_{-n}) = \{f \in C(\Gamma),\ f(\xi) = f(\zeta) \text{ if } \|\xi - \zeta\| < -n\}$. Denote by $\widehat{G}_n$ the linear operator on $\widehat{C}(\Gamma_{-n})$ with $(\widehat{G}_n f)|_{\Gamma_{-n}} = G_n(f|_{\Gamma_{-n}})$. Note that for $k \le n$,

$$\widehat{G}_n|_{\widehat{C}(\Gamma_{-k})} = \widehat{G}_k. \tag{5.5}$$

By $d(\delta, \varepsilon) := 2^{\|\delta-\varepsilon\|}$ a metric is given on Γ that induces the product topology on Γ. Note that $\widehat{C}(\Gamma) := \bigcup_{n\in\mathbb{N}} \widehat{C}(\Gamma_{-n})$ is dense in $C(\Gamma)$.

PROPOSITION 5.2. *Let $\widehat{G}$ be the linear operator on $\widehat{C}(\Gamma)$ such that*

$$\widehat{G}|_{\widehat{C}(\Gamma_{-n})} = \widehat{G}_n. \tag{5.6}$$

The closure G of $\widehat{G}$ is a Markov generator. We denote by $\gamma(t)$ the random walk induced by G. By $\widetilde{\beta}(\delta_1, \ldots, \delta_m; t)$ we denote the corresponding system of instantaneously coalescing random walks started in $(\delta_1, \ldots, \delta_m)$.

PROOF. By (5.5), $\widehat{G}$ is well defined and has a dense domain. Hence G is a well defined (unique-valued) linear operator. Fix $\lambda > 0$. Since G_n is a Markov generator for each $n \in \mathbb{N}$ we have $\mathscr{R}(\lambda - \widehat{G}_n) = \widehat{C}(\Gamma_{-n})$. So the range of $\lambda - G$ is dense, $\mathscr{R}(\lambda - G) = \widehat{C}(\Gamma)$, and hence G is recognized as a Markov generator. [For a treatment of this point see Liggett (1985), Chapter I.] $\square$

We assume $(\gamma(t), \gamma_1(t), \gamma_2(t), \ldots))$ to be defined on one probability space such that

$$\gamma_n(t) = \gamma(t)\big|_{\Gamma_{-n}}. \tag{5.7}$$

Now it is immediate that

$$(\gamma_n(t)_{t\geq 0}) \to (\gamma(t)_{t\geq 0}) \quad \text{as } n \to \infty, \text{ uniformly and a.s. in } \mathscr{D}([0, \infty[). \tag{5.8}$$

LEMMA 5.3.

$$\widetilde{\beta}_n(\delta_1, \ldots, \delta_m; t) \to \widetilde{\beta}(\delta_1, \ldots, \delta_m; t) \quad \textit{as } n \to \infty, \textit{ in distribution } \forall\, t \geq 0.$$

PROOF. Let

$$\tau_n = \inf\{t \geq 0\colon\ \gamma_n(t) \equiv 0\}, \tag{5.9}$$

$$\tau = \inf\{t \geq 0\colon\ \gamma(t) \equiv 0\}. \tag{5.10}$$

Now by (5.8) and right continuity of paths,

$$\tau_n \uparrow \tau \quad \text{a.s.} \tag{5.11}$$

Since we can assume that the systems $\widetilde{\beta}, \widetilde{\beta}_1, \widetilde{\beta}_2, \ldots$ are coupled so that

$$\widetilde{\beta}_1 \geq \widetilde{\beta}_2 \geq \cdots \geq \widetilde{\beta} \tag{5.12}$$

and since τ has no atoms, a simple induction argument yields the conclusion. $\square$

5.2. *Proof of Theorems* 2(a) *and* 5(a). By compactness of the state space it suffices to show convergence of moments,

$$\mathbf{E} \prod_{m\in\mathbb{Z}} \big(\Theta_{n-m}(\mathbb{X}(ts_n))\big)^{\psi_m} \to M_\psi \quad \text{as } n \to \infty, \tag{5.13}$$

where $\psi \in \mathbb{N}_0^{\mathbb{Z}}$ is finite and M_ψ is some real number. The martingale property then follows easily by symmetry.

Thus let

$$\psi = \mathbb{1}_{l_1} + \cdots + \mathbb{1}_{l_r}, \qquad l_1 \leq \cdots \leq l_r \in \mathbb{Z}, \tag{5.14}$$

and denote

$$\Gamma(\psi) = \{\bar{\delta} = (\delta_1, \dots, \delta_r) \in \Gamma^r\colon \ \|\delta_i\| \le l_i\},$$
$$\Gamma_{-n}(\psi) = \Gamma(\psi) \cap (\Gamma_{-n})^r. \tag{5.15}$$

Then

$$\begin{aligned}\mathbf{E}\big[(\Theta_{n-\cdot}(\mathbb{X}(ts_n)))^{\psi}\big] &= \mathbf{E}\bigg[\prod_{j=1}^{r} \Theta_{n-l_j}(\mathbb{X}(ts_n))\bigg]\\ &= \bigg(\prod_{j=1}^{r} \#\Xi_{n-l_j}\bigg)^{-1} \mathbf{E}\bigg[\sum_{\|\xi_1\|\le n-l_1} \cdots \sum_{\|\xi_r\|\le n-l_r} x_{\xi_1}(t(Nc)^{n+1})\\ &\qquad \cdots x_{\xi_r}(t(Nc)^{n+1})\bigg].\end{aligned} \tag{5.16}$$

By the duality lemma and the comparison lemma this equals

$$\begin{aligned}&\bigg(\prod_{j=1}^{r} \#\Xi_{n-l_j}\bigg)^{-1} \mathbf{E}\bigg[\sum_{\|\xi_1\|\le n-l_1} \cdots \sum_{\|\xi_r\|\le n-l_r} \theta^{\#\tilde{\eta}_n^{\{\xi_1,\dots,\xi_r\}^*}(t(Nc)^{1+n})}\bigg] + \varepsilon(n)\\ &\quad = \int \mathbf{E}^{\bar{\delta}^*}\big[\theta^{\#\tilde{\beta}_n(t)}\big]\mu^r(d\bar{\delta}|\Gamma_{-n}(\psi)) + \varepsilon(n).\end{aligned} \tag{5.17}$$

By Lemma 5.3 this tends to

$$M_\psi := \int \mathbf{E}^{\bar{\delta}^*}\big[\theta^{\#\tilde{\beta}(t)}\big]\mu^r(d\bar{\delta}|\Gamma(\psi)). \tag{5.18}$$ □

5.3. *Proof of Theorems* 2(b) *and* 5(b). It suffices to show [recall τ from (5.10) and note that here d plays the role of m in Theorems 2 and 5]

$$\mathbf{P}^{\delta}(\tau < t) \to \begin{cases} 0, & \text{as } d = \|\delta\| \to \infty,\\ 1, & \text{as } d = \|\delta\| \to -\infty,\end{cases} \qquad \forall\, t > 0 \tag{5.19}$$

since then [recall M_ψ from (5.18)]

$$M_{m\cdot \mathbb{1}_d} \to \begin{cases} \theta^m, & \text{as } d \to \infty,\\ \theta, & \text{as } d \to -\infty.\end{cases}$$

A straightforward computation using (2.16) and abbreviating $v = Nc/(Nc-1) + 1/(N-1)$ yields

$$\begin{aligned}\mathbf{E}^d\, e^{-\lambda\tau} &= \lim_{n\to\infty} \mathbf{E}^d\, e^{-\lambda\tau_n}\\ &= \bigg(\sum_{m=d-1}^{\infty} \frac{N^{-m}}{\vartheta v(Nc)^{-m} + \lambda} - \frac{N}{N-1}\frac{N^{1-d}}{\vartheta v(Nc)^{1-d} + \lambda}\bigg)\bigg/\\ &\qquad \bigg(\sum_{m=-\infty}^{\infty} \frac{N^{-m}}{\vartheta v(Nc)^{-m} + \lambda}\bigg),\end{aligned} \tag{5.20}$$

whereas

$$
\begin{aligned}
\mathbf{E}^d\, e^{-\lambda\tau'} &= \lim_{n\to\infty} \mathbf{E}^d\, e^{-\lambda\tau'_n} \\
&= \Bigg(\sum_{m=d}^{0} \frac{N^{1-m}}{v(Nc-1)(Nc)^{-m} - 1 + \lambda} \\
&\qquad - \frac{N}{N-1} \frac{N^{1-d}}{v(Nc-1)(Nc)^{1-d} - 1 + \lambda} + \frac{N}{N-1}\frac{1}{\lambda} \Bigg) \Bigg/ \\
&\quad \Bigg(\sum_{m=-\infty}^{0} \frac{N^{1-m}}{v(Nc-1)(Nc)^{-m} - 1 + \lambda} + \frac{N}{N-1}\frac{1}{\lambda} \Bigg).
\end{aligned}
\tag{5.21}
$$

Now (5.19) follows from

$$
\lim_{d\to-\infty} \mathbf{E}^d\, e^{-\lambda\tau} = \lim_{d\to-\infty} \mathbf{E}^d\, e^{-\lambda\tau'} = 1 \qquad \forall\, \lambda < \infty \tag{5.22}
$$

and

$$
\lim_{d\to\infty} \mathbf{E}^d\, e^{-\lambda\tau} = 0 \qquad \forall\, \lambda > 0. \tag{5.23}
$$

□

5.4. *Proof of Theorems* 2(c) *and* 5(c). We let $N \to \infty$ and indicate quantities with a superscript N. Observe

$$
\mathbf{E}^d\, e^{-\lambda\tau^N} \sim \Bigg(\sum_{m=d-1}^{0} \frac{N^{-m}}{(Nc)^{-m} + \lambda} \Bigg) \Bigg/ \Bigg(\sum_{m=-\infty}^{0} \frac{N^{-m}}{(Nc)^{-m} + \lambda} \Bigg) \quad \text{as } N \to \infty, \tag{5.24}
$$

$$
\mathbf{E}^d\, e^{-\lambda\tau'^N} \sim \Bigg(\sum_{m=d-1}^{0} \frac{N^{-m}}{(Nc)^{-m} - 1 + \lambda} \Bigg) \Bigg/ \Bigg(\sum_{m=-\infty}^{0} \frac{N^{-m}}{(Nc)^{-m} - 1 + \lambda} \Bigg) \quad \text{as } N \to \infty, \tag{5.25}
$$

Thus

$$
\begin{aligned}
&\mathbf{E}^d \exp(-\lambda\tau^N (Nc)^{-a}) \\
&\quad \to \begin{cases} 1 - c^{d-a} + \dfrac{(c-1)c^{d-a-1}}{1 + \lambda/c}, & \text{if } d - a \le 0, \\ 0, & \text{if } d - a > 0, \end{cases} \quad \text{as } N \to \infty.
\end{aligned}
\tag{5.26}
$$

The same holds for τ'^N if $a < 0$, whereas if $a = 0$ and $d < 0$,

$$
\mathbf{E}^d\, e^{-\lambda\tau'^N} \to 1 - c^d + \frac{c^d}{1 + \lambda/(c-1)} \quad \text{as } N \to \infty. \tag{5.27}
$$

Denote by $\mathscr{E}(m)$ the exponential distribution with mean m. Then

$$
\begin{aligned}
&\mathscr{L}^d[\tau^N (Nc)^{-a}] \\
&\quad \Rightarrow \begin{cases} (1 - c^{d-a})\delta_0 + (c-1)c^{d-a-1}\mathscr{E}\left(\dfrac{1}{c}\right) & \\ \quad + c^{d-a-1}\delta_\infty, & \text{if } d \le a, \\ \delta_\infty, & \text{if } d > a, \end{cases} \quad \text{as } N \to \infty,
\end{aligned}
\tag{5.28}
$$

as well as in the finite case if $a < 0$. On the other hand, for $a = 0$ and $d < 0$,

$$\mathscr{L}^d[\tau'^N] \Rightarrow (1 - c^d)\delta_0 + c^d \mathscr{E}\left(\frac{1}{c-1}\right) \quad \text{as } N \to \infty. \tag{5.29}$$

Introduce the first exit times of $\Gamma(d) := \{\delta \in \Gamma\colon\ \|\delta\| \le d\}$:

$$\sigma_n^N := \inf\{t \ge 0\colon\ \gamma^N(t) \notin \Gamma(n)\}. \tag{5.30}$$

As in (3.14) we obtain

$$\mathscr{L}^d[\sigma_n^N (Nc)^{-n}] = \begin{cases} \mathscr{E}(1), & \text{if } d \le n, \\ \delta_0, & \text{if } d > n. \end{cases} \tag{5.31}$$

Thus

$$\begin{aligned} &\lim_{N\to\infty} \mathbf{P}^\delta \left[\|\gamma(t(Nc)^n)\| = 1 + n \right] \\ &\qquad = 1 - \lim_{N\to\infty} \mathbf{P}^\delta \left[\|\gamma(t(Nc)^n)\| \le n \right] = 1 - e^{-t} \quad \text{if } \|\delta\| \le n. \end{aligned} \tag{5.32}$$

By (5.28),

$$\mathbf{P}^d[\sigma_a^N \le \tau^N \le t(Nc)^a] \to 0 \quad \text{as } N \to \infty$$

and thus

$$\mathbf{P}^d[\tau^N \le \sigma^N] \to \frac{c-1}{c} \quad \text{as } N \to \infty.$$

Hence we get

$$\lim_{N\to\infty} \mathbf{P}^d \left[\tau^N \le t(Nc)^d \middle| \tau^N \le \sigma_d^N \right] = 1 - ce^{-ct}. \tag{5.33}$$

The picture is as follows: For large N a particle at level d jumps in time scale $(Nc)^d$ at rate 1 to level $d+1$. Before it succeeds in doing so, it attempts to hit the origin with rate $c - 1$ in this scale.

Now consider the coalescing random walks. For $\psi = \mathbb{1}_{d_1} + \cdots + \mathbb{1}_{d_r}$ as above let

$$\Delta^N(\psi) = \{(\delta_1, \ldots, \delta_r) \in \Gamma^r\colon\ \|\delta_i - \delta_j\| = d_i \wedge d_j\}.$$

All starting points for $\widetilde{\beta}^N(t)$ in $\Delta^N(\psi)$ are equivalent by symmetry, so we indicate quantities with a superscript ψ. Let further

$$L_d^N(t) = \#\{\delta \in \widetilde{\beta}^N(t)\colon\ \|\delta\| \le d\}, \tag{5.34}$$

$$U_d^N(t) = \#\{\delta \in \widetilde{\beta}^N(t)\colon\ \|\delta\| > d\}. \tag{5.35}$$

The same type of argument as in Section 4 now yields that the $\binom{L_{d_0}^N(t)}{2}$ pairs of particles of level d_1 coalesce asymptotically independently at rate $2(c-1)(Nc)^{-d_1}$. Independently of this, each of the $L_{d_1}^N$ particles of level d_1 jumps to level $d_1 + 1$ at rate $(Nc)^{-d_1}$.

The limiting behavior of this will be modeled by the following definition.

DEFINITION 5.4 (Death–escape process). Let (A_t, B_t) be the $\mathbb{N}\times\mathbb{N}$-valued Markov process with generator

$$\mathscr{G}((a_1,b_1),(a_2,b_2)) = \begin{cases} 2(c-1)\binom{a_1}{2}, & \text{if } a_2 = a_1 - 1,\ b_2 = b_1, \\ a_1, & \text{if } a_2 = a_1 - 1,\ b_2 = b_1 + 1, \\ -a_1 - 2(c-1)\binom{a_1}{2}, & \text{if } a_2 = a_1,\ b_2 = b_1, \end{cases} \tag{5.36}$$

and let $G_t(m) = A_t + B_t$ if $(A_0, B_0) = (m, 0)$.

Particles in the first box (A) die with the same rate as they do in the pure death death process D_t of Definition 3.1. Here, however, they have a chance to escape to the second box (B) and remain there. Recall the definition of the Fisher–Wright diffusion X_t^θ with drift toward θ in (1.17). One easily checks the following duality relation.

LEMMA 5.5 (Duality).

$$\mathbf{E}^{(m,0)}[\theta^{G_t}] = \mathbf{E}[(X_t^\theta)^m]. \tag{5.37}$$

Let

$$\psi_d = \#\{\delta \in \widetilde{\beta}(0) \colon \|\delta\| = d\} \quad \text{and} \quad \psi_d^+ = \sum_{k>d} \psi^k.$$

Then

$$\mathscr{L}^\psi\big[L_{d_1}^N(t(Nc)^{d_1}), U_{d_0}^N(t(Nc)^{d_1})\big] \Rightarrow \mathscr{L}^\psi\big[(X_t, Y_t) | X_0 = \psi_{d_1},\ Y_0 = \psi_{d_1}^+\big] \quad \text{as } N \to \infty.$$

Thus

$$\mathscr{L}^\psi\big[\#\widetilde{\beta}^N(t(Nc)^{d_1})\big] \Rightarrow \mathscr{L}^\psi\big[\psi_{d_1}^+ + G_\infty(\psi_{d_1})\big] \quad \text{as } N \to \infty.$$

Iterating the argument and noting that $X_d^N(t(Nc)^{0.5+d}) \to 0$ as $N \to \infty$, we get

$$\mathscr{L}^\psi\big[\#\widetilde{\beta}^N\left((Nc)^{-0.5}\right)\big] \Rightarrow \mathscr{L}^\psi[\psi_{-1}^+ + G_\infty(\psi_{-1} + G_\infty(\psi_{-2} + \cdots + G_\infty(\psi_{d_1})\cdots))] \quad \text{as } N \to \infty.$$

Finally we get for the infinite system

$$\begin{aligned} &\mathscr{L}^\psi\big[\#\widetilde{\beta}^N(t)\big] \\ &\quad \Rightarrow \mathscr{L}^\psi[\psi_0^+ + G_t(\psi_0 + G_\infty(\psi_{-1} + G_\infty(\psi_{-2} + \cdots + G_\infty(\psi_{d_1})\cdots)))] \\ &\qquad \text{as } N \to \infty. \end{aligned} \tag{5.38}$$

In the last step the finite system differs from the infinite system since $\sigma_0 \equiv \infty$ is in the former and thus by (5.29) particles coalesce at rate $c-1$. Let $G_1' = D_{2(c-1)t}$. With this, (5.38) transforms to

$$\begin{aligned}&\mathscr{L}^{\psi}[\#\widetilde{\beta}'^N(t)]\\ \text{(5.39)}\qquad &\Rightarrow \mathscr{L}^{\psi}[\psi_0^+ + G_t'(\psi_0 + G_\infty(\psi_{-1} + G_\infty(\psi_{-2} + \cdots + G_\infty(\psi_{d_1})\cdots)))]\\ &\qquad\qquad \text{as } N\to\infty.\end{aligned}$$

Denote by $q_t(\psi, m)$ and $q_t'(\psi, m)$ the distribution

$$\begin{aligned}&q_t(\psi, m)\\ \text{(5.40)}\qquad &= \mathbf{P}^{\psi}[\psi_0^+ + G_t(\psi_0 + G_\infty(\psi_{-1} + G_\infty(\psi_{-2} + \cdots + G_\infty(\psi_{d_1})\cdots))) = m]\end{aligned}$$

in (5.38) and (5.39), respectively, and observe

$$(\mu^N)^r(\Delta(\psi)|\Gamma(\psi)) \to 1 \quad \text{as } N\to\infty. \tag{5.41}$$

Hence

$$M_\psi^N \to \sum_m q_t(\psi, m)\theta^m \quad \text{as } N\to\infty \tag{5.42}$$

and

$$M_\psi'^N \to \sum_m q_t'(\psi, m)\theta^m \quad \text{as } N\to\infty. \tag{5.43}$$

By the duality Lemma 5.5 the mixed moments of the Markov chains (Z_m^t) and $(\widetilde{Z}_m^t)$ defined in (1.20) and (1.28) are given by the right-hand sides of (5.42) and (5.43). Since $[0,1]^{\mathbb{Z}}$ is compact, the convergence of the mixed moments in (5.42) and (5.43) yields the assertions of Theorems 2(c) and 5(c). $\square$

Acknowledgment. It is a pleasure to thank Professor Andreas Greven for his valuable help in preparing this manuscript.

REFERENCES

ARRATIA, R. (1982). Coalescing Brownian motion and the voter model on $\mathbb{Z}$. Unpublished manuscript.

BAILLON, J.-B., CLÉMENT, PH., GREVEN, A. and DEN HOLLANDER, F. (1995). On the attracting orbit of a non-linear transformation arising from renormalization of hierarchically interacting diffusions. Part I: the compact case. *Canad. J. Math.* **47** 3–27.

BRAMSON, M. and GRIFFEATH, D. (1980). Clustering and dispersion rates for some interacting particle systems on $\mathbb{Z}$. *Ann. Probab.* **8** 183–213.

COX, J. T. (1989). Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$. *Ann. Probab.* **17** 1333–1366.

COX, J. T. and DURRETT, R. (1995). Hybrid zones and voter model interfaces. *Bernoulli* **1** 343–370.

COX, J. T., FLEISCHMANN, K. and GREVEN, A. (1996). Comparison of interacting diffusions and applications to their ergodic theory. *Probab. Theory Related Fields*. To appear.

COX, J. T. and GREVEN, A. (1990). On the long term behavior of some finite particle systems. *Probab. Theory Related Fields* **85** 195–237.

COX, J. T. and GREVEN, A. (1991). On the long term behavior of finite particle systems: a critical dimension example. In *Random Walks, Brownian Motion and Interacting Particle Systems: A Festschrift in Honor of Frank Spitzer* (R. Durrett and H. Kesten, eds.). Birkhäuser, Boston.

COX, J. T. and GREVEN, A. (1994a). Ergodic theorems for infinite systems of locally interacting diffusions. *Ann. Probab.* **22** 833–853.

COX, J. T. and GREVEN, A. (1994b). The finite systems scheme: an abstract theorem and a new example. *CRM Proceedings and Lecture Notes* **5** 55–67.

COX, J. T., GREVEN, A. and SHIGA, T. (1994). Finite and infinite systems of interacting diffusions. *Probab. Theory Related Fields.* To appear.

COX, J. T. and GRIFFEATH, D. (1986). Diffusive clustering in the two dimensional voter model. *Ann. Probab.* **14** 347–370.

DAWSON, D. A. and GÄRTNER, J. (1988). Long time behavior of interacting diffusions. In *Stochastic Calculus in Application: Proceedings of the Cambridge Symposium, 1987* (J. R. Norris, ed.). 29–54, Longman, Harlow, UK.

DAWSON, D. A. and GREVEN, A. (1993a). *Multiple Time Scale Analysis of Hierarchically Interacting Systems. A Festschrift to Honor G. Kallianpur* 41–50. Springer, New York.

DAWSON, D. A. and GREVEN, A. (1993b). Multiple time scale analysis of interacting diffusions. *Probab. Theory Related Fields* **95** 467–508.

DAWSON, D. A., GREVEN, A. and VAILLANCOURT, J. (1995). Equilibria and quasiequilibria for infinite systems of interacting Fleming–Viot processes. *Trans. Amer. Math. Soc.* **347** 2277–2361.

DURRETT, R. and SCHONMANN, R. H. (1988). The contact process on a finite set, II. *Ann. Probab.* **16** 1570–1583.

ERDÖS, P. and TAYLOR, S. J. (1960). Some problems concerning the structure of random walk paths. *Acta Math. Acad. Sci. Hungar.* **11** 137–162.

ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes, Characterization and Convergence*. Wiley, New York.

FLEISCHMANN, K. and GREVEN, A. (1994). Diffusive clustering in an infinite system of hierarchically interacting diffusions. *Probab. Theory Related Fields* **98** 517–566.

KEMENY, J. G., SNELL, J. S. and KNAPP, A. W. (1976). *Denumerable Markov chains*, 2nd ed. Springer, New York.

LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.

SAWYER, S. (1976). Results for the stepping stone model for migration in population genetics. *Ann. Probab.* **4** 699–728.

SHIGA, T. (1980). An interacting system in population genetics. *J. Math. Kyoto Univ.* **20** 213–242.

SHIGA, T. and SHIMIZU, A. (1980). Infinite dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.* **20** 395–416.

SPITZER, F. (1964). *Principles of Random Walk.* Van Nostrand, Princeton, NJ.

TRIBE, R. (1993). The long term behavior of a stochastic PDE. Preprint 74, Berlin.

YAMADA, T. and WATANABE, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11** 155–167.

MATHEMATISCHES INSTITUT
UNIVERSITÄT ERLANGEN-NÜRNBERG
BISMARCKSTRASSE $1\frac{1}{2}$
91054 ERLANGEN
GERMANY
E-mail: klenke@mi.uni-erlangen.de