Diffusive Clustering

Den Naturwissenschaftlichen Fakultäten
der Friedrich-Alexander-Universität
Erlangen-Nürnberg
zur

Erlangung des Doktorgrades

vorgelegt von

Achim Klenke

aus der Freien Hansestadt Bremen

Als Dissertation genehmigt von den Naturwissenschaftlichen Fakultäten der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung: 21. Februar 1996

Vorsitzender der Prüfungskommission:	Professor Dr. D. Kölzow
Erstberichterstatter:	Professor Dr. A. Greven
Zweitberichterstatter:	Professor Dr. D. Dawson (Carleton University, Ottawa)
Drittberichterstatter:	Professor Dr. A. Wakolbinger (Universität Frankfurt)

Abstract

In many infinite interacting particle systems in $\mathbb{R}^d$ the ergodic behaviour is dimension dependent. In high dimensions these systems are stable (i.e. non-trivial equilibria exist) while so-called clustering occurs in low dimensions. Roughly speaking, we say that the clustering is *diffusive* if the clusters (i.e. regions in which the system is "near" a constant state) grow in time at a random order of magnitude.

The main goal of this work is to illustrate that diffusive clustering is a phenomenon occurring in a wide range of models of interacting infinite particle systems. To this end we investigate two models in great detail. One model is *(critically binary) branching Brownian motion* on $\mathbb{R}^d$ and its high density limit, the so-called *Dawson-Watanabe process*. The other model considered here is linearly interacting diffusions with state space $[0, 1]$ and indexed by the so-called *hierarchical group* Ξ. Here a certain parameter c in the interaction terms plays the role played by the dimension in the models with eucledian site space. The (countable Abelian) hierarchical group Ξ pays respect to the biological notion of different degrees of relationships between individuals.

We show that in the critical dimension, i.e. $d = 2$ resp. $c = 1$, diffusive clustering occurs in both models. Thus we demonstrate that the phenomenon of diffusive clustering is not intrinsic to properties such as compactness of the state space or eucledian geometry of the index set. It is rather the fact that in the critical dimension the *recurrent potential kernel* is a slowly varying function that leads to diffusive clustering. This point is studied in detail for the hierarchical group.

Thus for diffusive clustering we obtain a universality in the diffusive term, the index set, and the initial law. As the only essentials for the occurrence of diffusive clustering we will recognise certain potential theoretic properties of the migration mechanism.

The second aim of this work is to relate the behaviour of our models to their "finite version". These are defined on bounded subsets of the index set. In this context we give a description in the fashion of the "finite systems scheme" introduced by Cox and Greven (1990) for the (stable) high dimensional cases $d \geq 3$ resp. $c < 1$. The scheme will be modified to cope with the critical dimension cases $d = 2$ resp. $c = 1$. We can show that diffusive clustering occurs even in the finite version models and we give a precise qualitative description.

Contents

Figures

Introduction

Preface

This work consists of three parts, one of introductory nature and and two more technical parts. Part II deals with interacting diffusions on the hierarchical group while Part III treats the branching models on $\mathbb{R}^d$. The two latter parts are completely self-contained and can be read independently. The introduction is written so as to encourage the non specialists to read it (specialists may skip it). It gives an overview of the topics treated, illustrates the basic ideas and exhibits the results in their most simple form. It shall serve mainly as to put the results of Part II and III in perspective. Those who are apprehensive for missing hints such as “existence and uniqueness were shown by ...” are referred to Part II and III, as well as those who are looking for information about the history of the treated problems or related ones.

I would like to express my gratitude to my supervisor Prof. A. Greven for introducing me to the subject and for his many helpful comments and suggestions during the preparation of this doctoral thesis.

1 The models

We shall give a short description of the models considered in this thesis. More details can be found in the respective chapters. Certain inconsistencies in the notation have historical reasons. We will stick to those and hope that this does not lead to any confusions.

We shall mainly consider spatial models for population growth and population composition as they occur in mathematical models in biology. The models cover the two cases of unbounded resources and independent evolution of families and a case of fixed resources and constant population size.

1.1 Branching models

The following (possibly infinite) particle system $(\eta_t)_{t\geq 0}$ on $\mathbb{R}^d$ will be called *(critical binary) branching Brownian motion* (shorthand BBM).

- Each particle has a random life time distributed according to an exponential mean c^{-1} $(c > 0)$ random variable.
- During its life time each particle moves according to a d-dimensional standard Brownian motion.
- At the end of its life each particle disappears. It gives rise to a random number of offspring located at the parents' position. The two possibilities, no offspring or two offspring, shall occur each with probability $\frac{1}{2}$.
- All random mechanisms are independent.

We consider η_t as a (random) measure on $\mathbb{R}^d$ by associating with each particle a unit mass at the respective position.

The *Dawson-Watanabe process* is the high density short life time limit of branching Brownian motion: For each $c > 0$ let $({}^c\eta_t)_{t\geq 0}$ be a BBM($\mathbb{R}^d$) with life time parameter c. Assume that the following limit of the initial configurations exists

$$\mathcal{L}[\frac{1}{c}{}^c\eta_0] \overset{c\to\infty}{\Longrightarrow} \mathcal{L}[\zeta_0].$$

(By "$\Longrightarrow$" we denote weak convergence and by "$\mathcal{L}$" the law of a random variable.) Then there exists a Feller process $(\zeta_t)_{t\geq 0}$ with state space in the Borel measures on $\mathbb{R}^d$ such that

$$\mathcal{L}[(\frac{1}{c}{}^c\eta_t)_{t\geq 0}] \overset{c\to\infty}{\Longrightarrow} \mathcal{L}[(\zeta_t)_{t\geq 0}]$$

(see Dawson (1993), Section 4.4 ff). This process (ζ_t) will be referred to as the *Dawson-Watanabe process* or *super Brownian motion* (shorthand SBM).

1.2 Voter Model

The voter model is that model for which diffusive clustering has first been studied (Cox and Griffeath (1986)). Also the relation of finite to infinite systems is well known in the case of the voter model (Cox (1989) and Cox and Greven (1990) and (1991)). Here the voter model will serve as an example to illustrate the basic ideas of diffusive clustering and finite systems.

The voter model $(\mathbb{V}(t))_{t\geq 0} = (v_i(t),\ i \in S)_{t\geq 0}$ is a continuous time model with site space $S = \mathbb{Z}^d$. Each site $i \in S$ is occupied by a person (the voter) capable of one opinion v_i in $\Sigma = \{0, 1\}$. Each person changes its opinion at rate equal to $\frac{1}{2d}$ times the number of (nearest) neighbours being in disagreement. This is, the person at site $i \in S$ changes its opinion $v_i(t)$ to $1 - v_i(t)$ at rate

$$\frac{1}{2d} \sum_{j\in S, |j-i|=1} |v_i(t) - v_j(t)|.$$

To be somewhat more general we let $a(i, j)$ the transition kernel of a symmetric random walk on a countable Abelian group S. We define the rate at which a person at i changes its opinion to be

$$\sum_{j\in S} a(i, j)|v_i(t) - v_j(t)|.$$

We will refer to this model as the *voter model on S with interaction kernel $a(i, j)$*. The special case of a Bernoulli random walk on $S = \mathbb{Z}^d$ is that of our nearest neighbour interaction defined above.

1.3 Interacting Diffusions

Consider a large population of individuals labelled by $1, \dots, n$. Each individual has genotype either A or B, say. Pairs of individuals interact in the following way. For each ordered pair (i, j) of individuals at rate 1 j changes its genotype to that of i. We may interpret this as "individual j dies at rate n and is replaced by an offspring of one of the remaining $n - 1$ individuals chosen at random". This is the so-called *Moran model*.

Of course, this model can also be considered as a (speeded up) voter model on $\{1, \dots, n\}$ without spatial structure, i.e. with $a(i, j) = 1/n\ \forall\, i, j \in \{1, \dots, n\}$.

As $n \to \infty$ the process of empirical frequencies converges to the so-called *Fisher-Wright diffusion* $(Y_t)_{t\geq 0}$. This is the diffusion process on $[0, 1]$ with generator

$$\frac{1}{2} g(x) \frac{\partial^2}{(\partial x)^2},$$

where $g(x) = x(1 - x)$ is called the *diffusion coefficient*. Assume now that at each site $i \in S$ there is located one (large) colony with empirical frequency $x_i(t)$. We will allow each colony to re-sample according to the procedure described above. In addition we will impose an interaction by allowing a migration between the colonies. The strength of migration shall be described by the kernel $a(i, j)$ of a random walk on S. Thus our system $\mathbb{X}(t) = (x_i(t),\ i \in S)$ of interacting diffusions is the Markov process on $[0, 1]^S$ with generator

$$\sum_{i,j\in S} a(i, j)\ (x_j - x_i) \frac{\partial}{\partial x_i} + \sum_{i\in S} \frac{1}{2} g(x_i) \frac{\partial^2}{(\partial x_i)^2}.$$

Here the diffusion coefficient $g(x)$ may be allowed to be somewhat more general but for simplicity we will defer this point to the more detailed description in Part II.

We will be particularly interested in $S = \Xi$ to be the *hierarchical group* defined below. Henceforth we will always assume $\mathbb{X}(t)$ to be defined on Ξ. Thus we refer to $\mathbb{X}(t)$ as the system of *hierarchically interacting diffusions*. The hierarchical group Ξ is defined by

$$\Xi := \{\xi = (\xi_m)_{m\in\mathbb{N}} : \xi_m \in \{0, \dots, N-1\}, \xi_m \neq 0 \text{ only for finitely many } m\}$$

with addition component wise modulo N ($N = 2, 3\dots$ is some fixed parameter) and distance $\|\xi\| := \max\{k : \xi_k \neq 0\} \vee 0$. For $n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ we denote by Ξ_n the finite subgroup

$$\Xi_n := \{\xi \in \Xi : \|\xi\| \leq n\}.$$

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

$$r_k := a(\xi, \zeta) R_k \qquad \text{with} \qquad R_k := \#\{\xi \in \Xi : \|\xi\| = k\} = (N - 1 + \mathbb{1}_0(k)) N^{k-1}. \tag{1}$$

In particular we will be concerned with the geometrical kernels a_c defined by $r_k = \vartheta c^k$, $k = 0, 1, \dots$ (ϑ a normalising constant). These are known to be recurrent iff $c \geq 1$.

The idea is that the colonies are organised 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 etc. 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 etc. 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. The flow spreads uniformly on the sites of degree k relatives ($k = 1, 2, \dots$).

2 Basic Ergodic Theory

Voter Model and Interacting Diffusions

Consider the voter model $(\mathbb{V}(t))$ in $\mathbb{Z}^d$, started in the product measure $\mathcal{L}[\mathbb{V}(0)] = \pi_\theta$ with parameter θ for some $\theta \in]0, 1[$. This is, all components are independent and have intensity

$$\theta = \mathbf{E}^{\pi_\theta}[v_i(0)] \qquad \forall\, i \in \mathbb{Z}^d.$$

A natural question to ask is whether for large times the voters are locally in consensus or not. This is, we ask if for finite $A \subset \mathbb{Z}^d$

$$\mathbf{P}^{\pi_\theta}[v_i(t) = 1, \forall\, i \in A] + \mathbf{P}^{\pi_\theta}[v_i(t) = 0, \forall\, i \in A] \overset{t\to\infty}{\longrightarrow} 1$$

holds. It is easily shown (see e.g. Liggett (1985)) that consensus will be obtained iff the transition kernel $a(i, j)$ generates a recurrent random walk and that the following holds

$$\mathcal{L}^{\pi_\theta}[\mathbb{V}(t)] \overset{t\to\infty}{\Longrightarrow} (1 - \theta)\delta_{\mathbf{0}} + \theta \delta_{\mathbf{1}}. \tag{2}$$

Here $\delta_{\mathbf{0}}$ and $\delta_{\mathbf{1}}$ are the Dirac measures on the configurations $\mathbf{0}, \mathbf{1} \in \{0, 1\}^{\mathbb{Z}^d}$ with all components 0 or 1. (We assume all product spaces to be equipped with the Tychonov topology. Hence weak convergence amounts to a local statement.) The behaviour is called *clustering* because there are growing clusters of voters in consensus.

On the other hand, if $a(i,j)$ generates a transient random walk, then there exists a family $(\nu_\theta,\ \theta \in [0,1])$ of invariant (under the dynamics) and shift ergodic measures with intensity $\int x_0 \nu_\theta(dx) = \theta$ such that

$$\mathcal{L}^{\pi_\theta}[\mathbb{V}(t)] \overset{t\to\infty}{\Longrightarrow} \nu_\theta. \tag{3}$$

This existence of non-trivial equilibrium states in the latter case is referred to as *stability* of the particle system. For the nearest neighbour interaction stability occurs iff $d \geq 3$. In this case there exists for each $\theta \in]0,1[$ a positive (bounded away from 0) probability of local disconsensus, $\liminf_{t\to\infty} \mathbf{P}^{\pi_\theta}[v_i(t) \neq v_j(t)] > 0 \ \forall\ i \neq j$.

In fact, it is not important that we have defined the voter model on the state space $S = \mathbb{Z}^d$. Any countable Abelian group will do it. The only point of the dichotomy between clustering and stability is that of recurrence or transience of $a(i,j)$.

Cox and Greven (1994a) have shown that this is true also for our systems $(\mathbb{X}(t))$ of linearly interacting diffusions with state space $[0,1]$. This is, (2) holds for $(\mathbb{X}(t))$ if $a(i,j)$ is recurrent. If $a(i,j)$ is transient then (3) holds with an, of course, different family $(\nu_\rho,\ \rho \in [0,1])$.

Branching Models

In order to discuss the long time behaviour of the branching models we have to talk about local *extinction* rather than consensus. We start with (η_t) BBM($\mathbb{R}^d$). It is well known that a family generated by one particle at time $t = 0$ eventually dies out. This is due to the criticality of the branching mechanism. So we may ask whether $\delta_{\mathbf{0}}$ (the Dirac measure on the empty configuration $\mathbf{0}$) is the only invariant measure for (η_t). Indeed Bramson et al. (1993) and (1995) have shown that this is true iff $d \leq 2$. Moreover in this case for initial configurations with asymptotic finite intensity (i.e. $\limsup_{R\to\infty} R^{-d}\mathbf{E}[\eta_0([-R,R]^d)] < \infty$) we have

$$\mathcal{L}[\eta_t] \overset{t\to\infty}{\Longrightarrow} \delta_{\mathbf{0}}. \tag{4}$$

On the other hand the branching system is *stable* if $d \geq 3$. This is there exists a one parameter convolution semigroup $(\nu_\rho,\ \rho \geq 0)$ of invariant (random) measures on $\mathbb{R}^d$ with intensity $\rho = |A|^{-1}\int \eta(A)\nu_\rho(d\eta)$. Moreover if we start in $\mathcal{L}[\eta_0] = \mathcal{H}(\rho) :=$ Poisson point process with intensity ρ then

$$\mathcal{L}^{\mathcal{H}(\rho)}[\eta_t] \overset{t\to\infty}{\Longrightarrow} \nu_\rho. \tag{5}$$

This behaviour is also called *persistence* since the intensity ρ is preserved in the limit.

As can be expected from the construction SBM shows the same ergodic behaviour as BBM with stability / persistence if $d \geq 3$ and extinction if $d \leq 2$. In particular (5) holds for SBM (for different ν_ρ, of course) if we let $\eta_0 = \rho \cdot \lambda$ a.s. (λ the d-dimensional Lebesgue measure).

Since BBM and SBM behave in a very similar manner in the sequel we will formulate most results simultaneously. For this purpose we denote by $(\psi_t)_{t\geq 0}$ either BBM or SBM. $M(\rho)$ will be $\mathcal{H}(\rho)$ resp. $\rho \cdot \lambda$. Thus, repeating the above statement, (ψ_t) becomes extinct if $d \leq 2$:

$$\mathcal{L}[\psi_t] \overset{t\to\infty}{\Longrightarrow} \delta_{\mathbf{0}}.$$

For each $d \geq 3$ there exists a family $(\nu_\rho,\, \rho \geq 0)$ of ergodic invariant (random) measures on $\mathbb{R}^d$ with intensity ρ and such that

$$\mathcal{L}^{M(\rho)}[\psi_t] \overset{t\to\infty}{\Longrightarrow} \nu_\rho \qquad \forall\, \rho \geq 0.$$

This is, (ψ_t) is persistent iff $d \geq 3$. The (ν_ρ) form a convolution semigroup,

$$\nu_{\rho+\sigma} = \nu_\rho * \nu_\sigma, \qquad \rho, \sigma \geq 0.$$

This latter property reflects the fact that particles were defined to move and branch independently.

By the criticality of the branching mechanism the "expected total population" remains constant. Thus local extinction goes along with the development of relatively slowly growing, but densely populated, areas. These will be called *clusters*.

The main point is now to investigate the behaviour of such clusters in more detail as a function of the diffusive part of the evolution, the structure of the site space (the index set for the components of the system) and properties of the migration mechanism.

3 Diffusive Clustering

We consider our interacting particle systems in low dimension, i.e. in the clustering regime. In order to better understand the clustering we may ask the following questions.

- How large is a cluster in terms of spatial extension? What is the law of growth?
- In the case of the branching models: How "high" is a cluster, i.e. how dense is the population?
- How old is a cluster?
- Will a particular site eventually be swallowed up by a cluster of a certain opinion resp. become depopulated?

In the case of diffusive clustering it will turn out that the answers to these questions are universal for the classes of models considered here. The occurrence of diffusive clustering does not rely on the details of the models but only on potential theoretic properties of the interaction kernel. The condition of slow variance of the recurrent potential kernel (explained below) ensures diffusive clustering. This condition also results in a close relationship to re-normalisation. In the frame-work of re-normalisation universality might seem to be more natural. For more on re-normalisation we refer to Dawson and Greven (1993b) and (1995) and Baillon et al. (1995).

3.1 Concepts of Clustering

The latter question can be reformulated in terms of the occupation time $T_t(i)$ of a site $i \in S$ resp. $T_t(B)$ ($B \subset \mathbb{R}^d$ measurable and bounded) for the branching models. $T_t(i)$ is defined by

$$T_t(i) = \int_0^t v_i(s)\, ds$$

for the voter models resp. $\int_0^t x_i(s)ds$ for interacting diffusions. For (ψ_t) either BBM or SBM we let

$$T_t(B) = \int_0^t \psi_s(B)\, ds.$$

So we reformulate our question to: Does the limit $T(i) := \lim_{t\to\infty} t^{-1}T_t(i)$, resp. $T(B) := \lim_{t\to\infty} t^{-1}T_t(B)$ exists? If it exists, are the random variables $T(i)$ resp. $T(B)$ trivial in the sense that they are concentrated on $\{0,1\}$ resp. 0? Or does the limit exist and is almost surely constant with a non-trivial value? This latter alternative implies of course that the state of a particular site changes infinitely often.

This question has been dealt with by Cox and Griffeath (1983) for the voter model (on $\mathbb{Z}^d$) and (1985) for BBM. Iscoe (1986), Thm. 4.3, treats the question for SBM. We study interacting Fisher-Wright diffusions $(\mathbb{X}(t))$ on Ξ in Part II, Section 6. It has been shown for the low dimension ($d = 1$, resp. $c > 1$ for the geometrical kernel a_c on Ξ, defined above) that $T(A) \equiv 0$ for (ψ_t) and that there is no law of large number neither for the voter model nor for interacting diffusions on Ξ. On the other hand, $T(A)$ is non degenerate if $d = 2$. For the voter model and interacting diffusions on Ξ in the critical dimension ($d = 2$ resp. $c = 1$) we have a law of large number $\mathcal{L}^{\pi_\theta}[T(0)] = \delta_\theta$. Indeed, e.g. for interacting diffusions on Ξ we have $\mathbf{E}^{\pi_\theta}[t^{-1}T_t(0)] = \theta$ and our Proposition II.6.1 claims

$$\mathrm{Var}^{\pi_\theta}[t^{-1}T_t(0)] \overset{t\to\infty}{\longrightarrow} 2\theta(1-\theta)(1-2^{-\log c/\log cN}) \qquad \text{for } c \geq 1.$$

The question concerning the age of the clusters has been raised only recently and is studied in detail in Fleischmann and Greven (1995).

The first of the four questions raises the the problem to define the clusters more precisely. In the case of the one dimensional voter model this is easy. Here a cluster is a maximal interval $\{m, \ldots, n\} \subset \mathbb{Z}$ containing only 0's or 1's. For the one dimensional interacting diffusions we have the notion of ε-clusters. These are the maximal intervals with all components in $[0, \varepsilon[$ or $]1-\varepsilon, 1]$. In higher dimensions and for Ξ there are no satisfying precise definitions of clusters.

First we consider the voter model and interacting diffusions. Here we circumvent this problem by investigating the following two, closely related, quantities:

- The correlation function of two components $x_i(t)$ and $x_j(t)$. The more $x_i(t)$ and $x_j(t)$ are correlated the more we will be willing to say that i and j are situated in the same cluster.

- The *block average* of a finite set $A \subset S$

$$\rho_A(t) = \frac{1}{\#A} \sum_{i\in A} x_i(t).$$

  The closer $\rho_A(t)$ is to either 0 or 1 we will be willing to say that A is contained in a cluster.

For the branching models only slight modifications of these quantities are necessary.

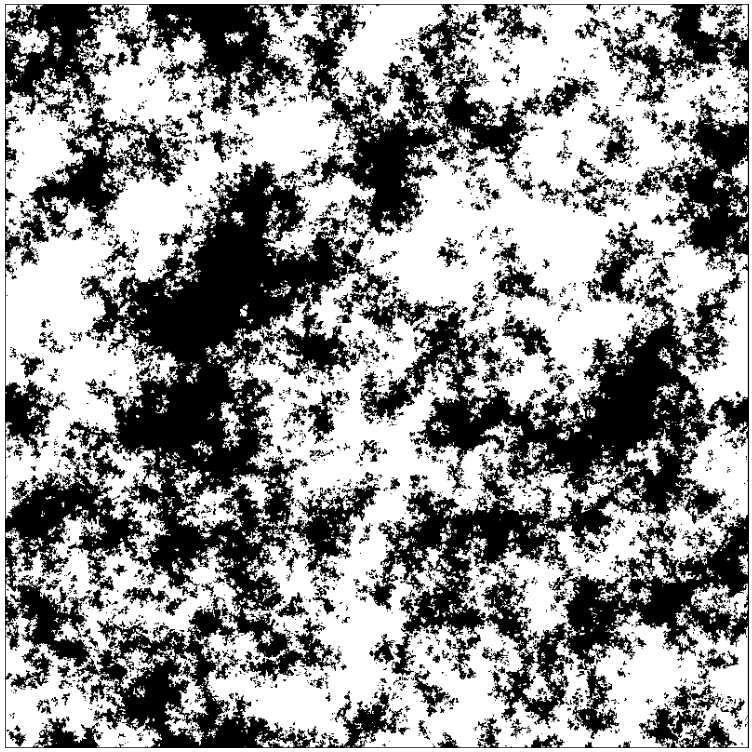

Figure 1. Simulation of the voter model on a 800×800 *grid at time* $t = 100,000$ *started with intensity* $\theta = \frac{1}{2}$.

3.2 Diffusive Clustering for the Voter Model

We illustrate the concepts with the example of the voter model on $\mathbb{Z}^2$. The results of this subsection are taken from Cox and Griffeath (1986).

First consider both cases $d = 1$ and $d = 2$. In order to give an indication of the speed at which clusters grow we introduce a scaling function $f_\alpha(t) \uparrow \infty$, as $t \to \infty$, and a decreasing function $h(\alpha)$, $\alpha \in I$ defined on an interval $I \subset \mathbb{R}$. We hope to find $f_\alpha(t)$ and $h(\alpha)$ such that

$$\mathrm{Cov}^{\pi_\theta}[v_i(t), v_j(t)] \sim \theta(1-\theta)\cdot h(\alpha),\ \text{as } t \to \infty, \tag{6}$$

if i and j are points with increasing distance such that $\mathrm{dist}(i,j) \sim f_\alpha(t)$ as $t \to \infty$. Here $h(\alpha)$ measures "how much" or "how probably" the considered sites are in one cluster. Finally the scaling function $f_\alpha(t)$ describes the speed of growth of an area of sites being correlated more than $\theta(1-\theta)h(\alpha)$.

Consider the *rescaled* or *thinned out* system ${}^\alpha\mathbb{V}(t) = ({}^\alpha v_z(t), z \in \mathbb{R}^d)$, defined by

$${}^\alpha v_z(t) = v_{([zf_\alpha(t)])}(t), \qquad z \in \mathbb{R}^d$$

($[zf_\alpha(t)]$ is the nearest lattice point to $zf_\alpha(t)$ with some convention in case of ties). We might even hope that $({}^\alpha v_z(t),\ z \in \mathbb{R}^d)$ converges to a limit field ${}^\alpha\mathbb{V}(\infty) = ({}^\alpha v_z(\infty),\ z \in \mathbb{R}^d)$ as $t \to \infty$.

Both hopes are in fact justified. In dimension $d = 1$ it turns out that $f_\alpha(t) = \alpha\sqrt{t}$ is the right scale where

$$h(\alpha) = \frac{1}{\pi} \int_0^1 \frac{e^{-\alpha^2/4s}}{\sqrt{s(1-s)}} \, ds.$$

In fact even more is true. Arratia (1982) shows that the process of edges between clusters of 0's and clusters of 1's in $(v_{[zt^{1/2}]}(st),\ z \in \mathbb{R})_{s \geq 0}$ converges as $t \to \infty$ to a system of annihilating Brownian motions.

The same scale appears if we investigate the block averages $\rho_{A_t}(t)$, where $A_t = \{i \in \mathbb{Z}^2,\ |i| \leq \sqrt{t}\}$. In fact, the limit of $\rho_{A_t}(t)$ as $t \to \infty$ exists and is non degenerate.

Thus with some right we may say that in the one dimensional voter model the size of clusters is of order $\sqrt{t}$.

Consider now the voter model on $\mathbb{Z}^2$. Here we have that (6) holds with $I = [0,1],\ f_\alpha(t) = t^{\alpha/2}$ and $h(\alpha) = 1 - \alpha$. The same scale function $f_\alpha(t) = t^{\alpha/2}$ is crucial for the description of the block averages. Let $A_t = \{i \in \mathbb{Z}^2,\ |i| \leq t^{\alpha/2}\}$. Then

$$\mathcal{L}^{\pi_\theta}[\rho_{A_t}(t)] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}^\theta[Y_{\hat{\alpha}}],$$

where $\hat{\alpha} := -\log \alpha$ and $(Y_s)_{s \geq 0}$ is a Fisher-Wright diffusion. The convergence is known to hold in the sense of finite dimensional distributions when regarded as processes in α. The limit field ${}^\alpha\mathbb{V}(\infty)$ exists, is exchangeable, and can be described explicitly by means of its de Finetti representation

$$\mathcal{L}^{\pi_\theta}[{}^\alpha\mathbb{V}(\infty)] = \int_0^1 \mathbf{P}^\theta[Y_{\hat{\alpha}} \in d\rho]\pi_\rho,$$

where the latter π_ρ means the product measure on $\{0,1\}^{\mathbb{R}^2}$ with parameter ρ.

Thus the picture is somewhat different from the one dimensional case:

- Cluster growth is of order $t^{\alpha/2}$ for different $\alpha \in [0,1]$, i.e. the clusters grow at a random order of magnitude. (Note that this is qualitatively different from the $d = 1$ case. There clusters grow at a random *multiple* of the fixed scale $t^{1/2}$.) This is the property the word "diffusive" in the term "diffusive clustering" refers to.

- Going from large blocks ($\alpha = 1$) to small blocks ($\alpha = 0$) the block averages (in the limit $t \to \infty$) evolve as a *Markov* process.

- The limit field is exchangeable and can be described in terms of the same Markov process.

These three points are characteristic for diffusive clustering.

3.3 Diffusive Clustering for the Hierarchical Group

This subsection is devoted to show how the concepts of scaling, block averages etc. carry over to the non eucledian site space Ξ. Except for the notation nothing changes, however, by substituting interacting diffusions for the voter model. In fact, our Theorems II.1, II.2, II.4 and II.5 also hold for the voter model on Ξ. Theorem II.3 holds for the voter model with $V = 0$ in (1.29).

(i) Scaled systems

For a systematic treatment 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 $f : \mathbb{N}_0 \to \mathbb{N}_0, f(n) \uparrow \infty$ is called *space scale*. To keep time continuous we also introduce the "inverse" of (s_n)

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

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

Let the shift operators $S_k : \Xi \longrightarrow \Xi$, $k = 0, 1, 2, \ldots$, be defined by

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

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

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

(ii) Block averages

For $n \in \mathbb{N}$ let the nth block average be defined by

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

Fleischmann and Greven (1994) have studied the case of geometrical kernels a_c with $c = 1$. They show that the clustering is diffusive in the sense of the three criteria given in the last subsection. More precisely, they show for the following choice of scales

$$f_\alpha(n) = [\alpha n] \quad \text{and} \quad s_n = N^n$$

that the following holds

$$\mathcal{L}^{\pi_\theta}[\Theta_{f_\alpha(n(t))}(\mathbb{X}(t))] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}^\theta[Y_{\hat\alpha}] \tag{7}$$

and

$$\mathcal{L}^{\pi_\theta}[{}^{f_\alpha}(\mathbb{X}(t))] \overset{t\to\infty}{\Longrightarrow} \nu_\theta(\hat\alpha) := \int_0^1 \mathbf{P}^\theta[Y_{\hat\alpha} \in d\rho]\pi_\rho. \tag{8}$$

Note that the volume of a cluster described by $f_\alpha(n)$ is $N^{[\alpha n]}$. Thus we see that clusters grow indeed at random order of magnitude.

In Theorem II.1 we show that diffusive clustering occurs for the whole class of so-called *critical* kernels. These are defined by the property (which we explain in terms of potentials in a second) that (recall r_k from (1))

$$\log(k)[\log(r_k N^k) - \log(r_{k+1} N^{k+1})] \quad \text{is bounded.} \tag{9}$$

These kernels are recurrent. In particular a_1 is critical.

We will describe $f_\alpha(n)$ in terms of the recurrent potential kernel

$$A(\zeta, \xi) = \sum_{m=0}^\infty \left(a^{(m)}(\zeta,\zeta) - a^{(m)}(\zeta,\xi) \right).$$

Here $a^{(m)}(\zeta,\xi)$ denotes the m-step transition probabilities of the random walk generated by $a(\zeta,\xi)$. Let also
$$A(n) = \sup\{A(0,\xi),\ \xi \in \Xi_n\}.$$

Condition (9) guarantees that $A(n)$ is slowly varying as $n \to \infty$. This means that the large-scale properties of the model are in a sense slowly varying. Cluster growth turns out to become random at the scale at which the recurrent potential varies. Thus (9) assures cluster growth at a random order of magnitude.

In Theorem II.1 we show that (7) and (8) hold for $f_\alpha(n)$ defined such that

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

and with time scale $s_n = N^n A(n)$.

The scales of cluster growth are diffusive in the sense that $f_\alpha(n) - f_\beta(n) \stackrel{n\to\infty}{\longrightarrow} \infty$ for $\alpha > \beta$. However, we observe different sizes of clusters for different choices of $a(\cdot,\cdot)$:

$*$	small clusters	when	$\dfrac{f_\alpha(n)}{n} \stackrel{n\to\infty}{\longrightarrow} 0$	for $\alpha < 1$
$*$	medium clusters	when	$\lim_{n\to\infty} \dfrac{f_\alpha(n)}{n} \in]0,1[$	for $\alpha \in]0,1[$
$*$	large clusters	when	$\dfrac{f_\alpha(n)}{n} \stackrel{n\to\infty}{\longrightarrow} 1$	for $\alpha > 0$.

For instance these above cases can occur if we choose

$*$	$r_k = \vartheta k N^{-k}$	and	$f_\alpha(n) = n^\alpha$
$*$	$r_k = \vartheta N^{-k}$	and	$f_\alpha(n) = \alpha n$
$*$	$r_k = \vartheta k^{-\log k} N^{-k}$	and	$f_\alpha(n) = n\left(1 + \frac{\log \alpha}{2 \log n}\right)$.

The criticality of the kernel $a(\xi,\zeta)$ makes sure that $A(n)$ is a slowly varying function of the volume $\#\Xi_n$. As we see in Section II.2 this point seems to be crucial for diffusive clustering. Note also that for nearest neighbour interaction on $\mathbb{Z}^2$ and for Brownian motion on $\mathbb{R}^2$ the corresponding recurrent potential kernel grows on a logarithmic scale.

For geometrical kernels a_c with $c > 1$ the clusters are so large that it is suitable to take $f_\alpha(n) = n - \alpha$, $\alpha \in \mathbb{Z}$ but in this case in the sequel we will not stick to that notation. In this case a suitably modified version of (7) holds. Namely for fixed N the averages over the blocks of size $f_\alpha(n(t))$ (in the limit $t \to \infty$) form a discrete time martingale in α. While this martingale is not Markov for fixed N, we show that it is Markov in the re-normalisation limit $N \to \infty$ and we determine its transition law. This is the content of our Theorem II.2.

3.4 Branching Models

For the branching models we have to deal with the question of the height of a cluster. We start with some heuristics for BBM($\mathbb{R}^2$).

Consider at time t a particle located at, say, the origin. It is known that during its life time it has given rise to an offspring of approximately t particles. We also pretend that the time points s of branchings are distributed uniformly on $[0,t]$. Now that the transition density $p_{t-s}(x,y)$ of the Brownian motion on $\mathbb{R}^2$ is of order $\sim (t-s)^{-1}$ for $t-s$ large and $|x-y|$ small, we see that the expected number of particles in $[0,1]^2$ should be of order

$$\int_0^{t-1} \frac{1}{t-s}\, ds = \log t.$$

Fleischman (1978) has the more precise statement for (η_t) BBM($\mathbb{R}^2$)

$$\frac{\log t}{8\pi}\mathbf{P}^{M(1)}\left[\eta_t(B) > \frac{\log t}{8\pi}|B|x\right] \overset{t\to\infty}{\longrightarrow} e^{-x}, \quad x>0,\ B\in\mathcal{B}(\mathbb{R}^2).$$

Thus with probability $\sim \frac{8\pi}{\log t}$ we see a cluster. If we see a cluster then it is of size $\frac{\log t}{8\pi}$ times an exponential mean 1 random variable. So we have to do three things to describe clusters in these branching models:

- Make sure that we see a cluster.
- Scale down its height to a non-trivial size.
- Impose the spatial rescaling introduced in the last subsection.

To make sure that we see a cluster we could follow Fleischman and condition on this event. However, we prefer to take another approach. We start with more and more densely populated initial configurations, an approach often used in statistical physics models. This increase in density will be done so carefully that we do not loose too much information caused by a possible overlap of clusters. The blow-up of the initial configuration also serves to underline the similarities to other models of interacting particle systems or diffusions. The parallels are exhibited most clearly in this blow-up picture.

These considerations motivate the following definitions.

(1) Blow-up

At time $t>1$ we define

$$\widetilde{\psi}_t = \widetilde{\psi}_t^0 := \frac{8\pi}{\log t}\psi_t \tag{10}$$

with

$$\mathcal{L}[\psi_0] = \widetilde{M}(t) := M\left(\frac{\log t}{8\pi}\right). \tag{11}$$

(2) Spatial rescaling

For (ψ_t) BBM resp. SBM let $I=[0,1]$ resp. $I=]-\infty,1]$. For fixed $\alpha\in I$ we define $(\widetilde{\psi}_t^\alpha)$ by

$$\widetilde{\psi}_t^\alpha(B) = t^{-\alpha}\widetilde{\psi}_t(t^{\alpha/2}B). \tag{12}$$

As above we let $\widetilde{\psi}_t = \widetilde{\psi}_t^0$.

Now we can formulate the first result concerning the block averages. Our Theorem III.1 says that for fixed $\alpha\in I$

$$\mathcal{L}^{\widetilde{M}(t)}[\widetilde{\psi}_t^\alpha([0,1[^2)] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}^1[Z_{1-\alpha}].$$

Here and in the sequel $(Z_t)_{t\geq 0}$ is *Feller's continuous state branching diffusion.* This is the diffusion on $[0,\infty[$ with generator

$$x\frac{\partial^2}{(\partial x)^2}. \tag{13}$$

To formulate a finer result in the fashion of the rescaled limit field we have to introduce the following object: For $\tau > 0$ let $(Z_t^{\tau,1}),\dots,(Z_t^{\tau,n})$ be n (not independent) Feller diffusions with the following properties: $Z_t^{\tau,1} = Z_t^{\tau,2} = \dots = Z_t^{\tau,n}$ for all $t \leq \tau$. After time $t = \tau$ the diffusions evolve independently. Thus the common distribution at time $t \geq \tau$ is

$$\mathcal{L}^\rho\left[\left(Z_t^{\tau,k}\right)_{k=1,\dots,n}\right] = \int_0^\infty \mathbf{P}^\rho[Z_\tau \in d\rho'] \bigotimes_{k=1}^n \mathcal{L}^{\rho'}[Z_{t-\tau}], \tag{14}$$

where (Z_t) is an ordinary Feller diffusion.

Now we are able to describe the asymptotic dependence structure between various parts of the space, viewed in the $t^{\alpha/2}$-scale. Namely our Theorem III.2 says that for mutually distinct points $x_1,\dots,x_n \in \mathbb{R}^2$ the following holds

$$\mathcal{L}^{\widetilde{M}(t)}\left[(\widetilde{\psi}_t(t^{\alpha/2}x_k + [0,1[^2))_{k=1,\dots,n}\right] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}\left[\left(Z_1^{1-\alpha,k}\right)_{k=1,\dots,n}\right], \quad \alpha\in[0,1]$$
$$\mathcal{L}^{\widetilde{M}(t)}\left[(\widetilde{\psi}_t^\alpha([0,1[^2))_{\alpha\in I}\right] \underset{\text{fdd}}{\overset{t\to\infty}{\Longrightarrow}} \mathcal{L}^1\left[(Z_{1-\alpha})_{\alpha\in I}\right], \qquad B\in\mathcal{B}(\mathbb{R}^d),$$

(By "$\underset{\text{fdd}}{\overset{t\to\infty}{\Longrightarrow}}$" we denote weak convergence of the processes in the sense of their finite dimensional distributions).

Incidently, our Theorem III.2 gives an even more detailed description in terms of points being spaced at multiple scales (see Figure III.1 on page 69). For simplicity, however, we will not stress this point here.

Remarks

1. Note that this result is consistent with (8). In fact, we could replace in the definition of $(Z_t^{\tau,1}),\dots,(Z_t^{\tau,n})$ the Feller diffusions by Fisher-Wright diffusions to get $(Y_t^{\tau,1}),\dots,(Y_t^{\tau,n})$. With this notation (8) becomes

$$\mathcal{L}^\theta[(Y_{\hat{0}}^{\hat{\alpha},k})_{k=1,\dots,n}] = \mathcal{L}^\theta[(Y_{\infty}^{\hat{\alpha},k})_{k=1,\dots,n}] = \int_0^1 \mathbf{P}^\theta[Y_{\hat{\alpha}} \in d\rho]\pi_\rho^{(n)},$$

   where $\pi_\rho^{(n)}$ denotes the product measure on $\{0,1\}^n$ with parameter θ.

2. The clustering is recognised to be diffusive in the sense of the criteria given in Section 4.2.

4 Finite systems

4.1 Motivation

The theory of interacting particle systems investigates models coming from biology, chemistry and physics. For some reasons physical laboratories are not suited to contain, say, crystals of infinite size. Similar problems are known to biologists and chemists. This should be enough reason to study finite systems of interacting particle systems.

Most often however, it is analytically more convenient to study infinite systems. Thus the idea is to first analyse the infinite system and the relate its behaviour to the finite system. This has been suggested by Dobrushin (1971) in the context of spin flip systems.

Now imagine that for some reason an infinite particle system *is* an appropriate model for something. Nevertheless computer simulations of this system have to be restricted to finite versions of the model. So in order to be able to rely on computer simulations in this case we have to know how much and how the finite systems differ from the infinite one. (For more on this point see Durrett (1988).) This gives another justification to compare finite systems with infinite ones.

4.2 Concepts, History of the Subject

The raised problem has been attacked by various methods. These include

- Asymptotics of extinction and trapping times. These have been studies for the contact process by Griffeath (1981), Cassandro et al. (1984), Schonmann (1985), Durrett and Liu (1987), Durrett and Schonmann (1988) and Durrett et al. (1989). For the voter model this has been done by Donnelly and Welsh (1983) and Cox (1989).

- Asymptotics of the tunnelling times from meta stable states to thermodynamically stable states. See Dawson and Gärtner (1988).

In this work we will follow another approach, the so-called "finite systems scheme". This has been developed in Cox and Greven (1990), (1991), Cox, Greven and Shiga (1995) and Dawson and Greven (1993). Roughly speaking this scheme postulates the existence of a macroscopic system variable that dominates the behaviour of the finite system. Given the value θ of this variable the system should be in a state near the equilibrium of the infinite system corresponding to that value θ. The further investigation focuses on the behaviour of this macroscopic variable. Of course this scheme may work in this form only if we are in the stable case. For clustering models we will have to modify the scheme with respect to the rescaling ideas introduced above, since we do not have equilibrium states.

4.3 Interacting Diffusions

We want to explain the ideas of the finite systems scheme with the example of interacting diffusions on the hierarchical group. We start with the definition of the finite system.

Fix $n \in \mathbb{N}$ and let $\Xi_n = \{\xi \in \Xi,\ \|\xi\| \le n\}$ as above. Let $a(\cdot,\cdot)$ be the interaction kernel for the *infinite system* $(\mathbb{X}(t))$. We define the *finite system* (of hierarchically interacting diffusions) $(\mathbb{X}_n(t))$ on Ξ_n to be the Markov process on $[0,1]^{\Xi_n}$ with generator

$$\sum_{\xi \in \Xi_n} \left[\sum_{\zeta \in \Xi_n} a_n(\xi,\zeta)(x_{n,\zeta} - x_{n,\xi}) \right] \frac{\partial}{\partial x_\xi} + \sum_{\xi \in \Xi_n} \frac{1}{2} g(x_{n,\xi}) \frac{\partial^2}{(\partial x_\xi)^2} \qquad (\xi \in \Xi_n),$$

where

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

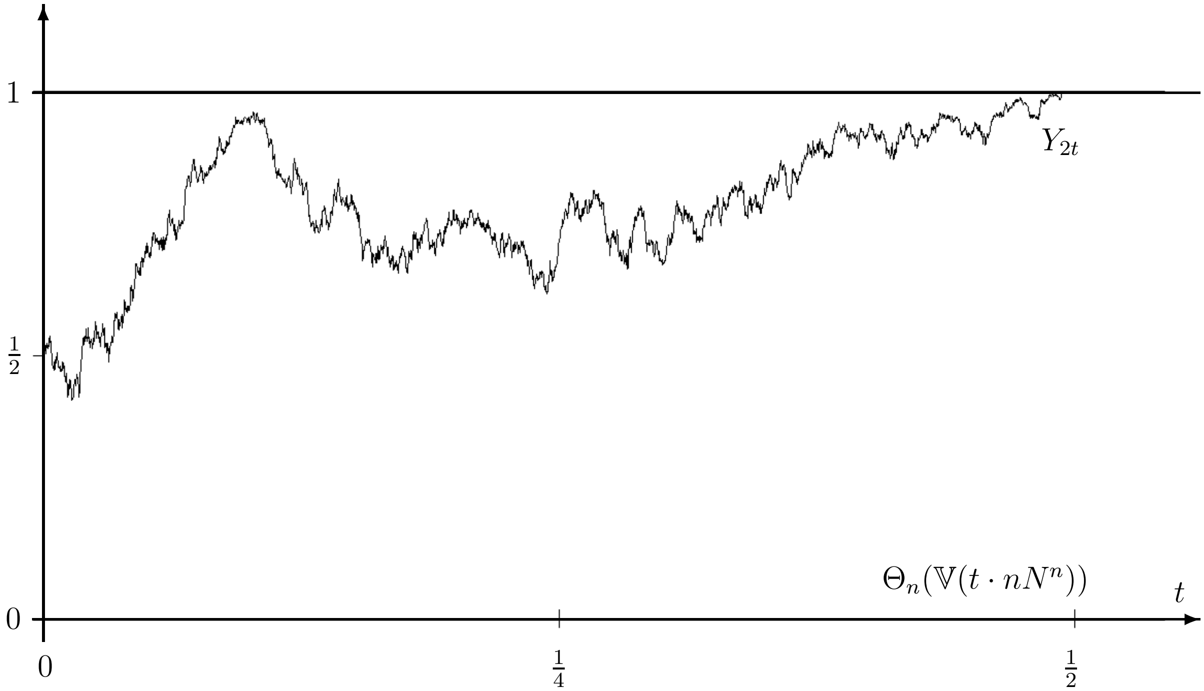

Figure 2. Block average $\Theta_n(\mathbb{V}(t \cdot nN^n))$ of a simulated voter model $\mathbb{V}$ on the finite hierarchical group Ξ_n (with $n = 10$ and $N = 2$) and a sample path of the Fisher-Wright-diffusion Y_{2t}. The initial configuration is $\mathcal{L}[\mathbb{V}(0)] = \pi_{1/2}$ resp. $Y_0 = \frac{1}{2}$.

Here a_n is the interaction kernel restricted to Ξ_n with periodic boundary conditions. Since a_n is recurrent, by the basic ergodic theorem for fixed $n \in \mathbb{N}$

$$\mathcal{L}^{\pi_\theta}[\mathbb{X}_n(t)] \overset{n\to\infty}{\Longrightarrow} (1-\theta)\delta_{\mathbf{0}} + \theta\delta_{\mathbf{1}}. \tag{15}$$

Assume henceforth that $a(\cdot,\cdot)$ is transient. Then we may expect that for $s'_n \uparrow \infty$ very slowly that the finite and the infinite systems agree, i.e. we expect

$$\mathcal{L}^{\pi_\theta}[\mathbb{X}_n(s'_n)] \overset{n\to\infty}{\Longrightarrow} \nu_\theta.$$

On the other hand, by (15), if we take $s''_n \uparrow \infty$ very fast then $\mathcal{L}^{\pi_\theta}[\mathbb{X}_n(s''_n)] \overset{n\to\infty}{\Longrightarrow} (1-\theta)\delta_{\mathbf{0}} + \theta\delta_{\mathbf{1}}$, i.e. the finiteness is exhibited.

Somewhere between (s'_n) and $s''_n)$ we expect to find a *time scale* (s_n) with a non-trivial limit of $\mathbb{X}_n(s_n)$.

Recall that $\Theta_n(\mathbb{X}_n(t)) = (\#\Xi_n)^{-1}\sum_{\xi\in\Xi_n} x_{n,\xi}(t)$ is the block average over Ξ_n. By the above discussion we should have

$$\mathcal{L}^{\pi_\theta}[\Theta_n(\mathbb{X}_n(s'_n))] \overset{n\to\infty}{\Longrightarrow} \delta_\theta = \mathcal{L}^\theta[Y_0]$$

(recall that ν_θ is shift ergodic) and

$$\mathcal{L}^{\pi_\theta}[\Theta_n(\mathbb{X}_n(s''_n))] \overset{n\to\infty}{\Longrightarrow} (1-\theta)\delta_0 + \theta\delta_1 = \mathcal{L}^\theta[Y_\infty].$$

Thus we would like to choose (s_n) such that the limit of the block averages is a non-trivial random variable. In fact, Theorem II.3 (a) says that $s_n = \text{const.} \cdot N^n$ (with the constant defined in (1.30)) is the right time scale. Moreover for $t > 0$

$$\mathcal{L}^{\pi_\theta}[\Theta_n(\mathbb{X}_n(s'_n))] \overset{n\to\infty}{\Longrightarrow} \mathcal{L}^\theta[Y_{2t}],$$

where as usual (Y_t) is a Fisher-Wright diffusion (see Figure 2).

Now we explain the main point of the finite systems scheme. Increasing t we see that the block average fluctuates at times of order s_n. On the other hand at times of order $\ll s_n$ our finite systems behave almost like the the infinite one. Thus we might expect that given the value $\Theta_n = \rho$ our finite system $(\mathbb{X}_n(t))$ *relaxes* towards a state near the equilibrium ν_ρ for $(\mathbb{X}(t))$. We may also assume that this relaxation takes place faster than the fluctuations of the block averages. Indeed, we can show (Theorem II.3 (b)) that an integral statement of this holds. Namely

$$\mathcal{L}^{\pi_\theta}[\mathbb{X}_n(ts_n)] \overset{n\to\infty}{\Longrightarrow} \int_0^1 \mathbf{P}^\theta[Y_{2t} \in d\rho]\, \nu_\rho.$$

So far we have talked about the stable case in which $a(\cdot,\cdot)$ is transient. Let us now turn to the case in which $a(\cdot,\cdot)$ is critically recurrent, i.e. diffusive clustering occurs.

From a systematic point of view the results are quite similar to those of the stable case. In order to see this recall $A(n)$, $f(n)$, $n(t)$ and $s(n) = A(n) \cdot N^n$ from Section 4.3. We define the rescaled finite system $({}^f\mathbb{X}_n(t))$ by

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

It turns out that $s_n = A(n) \cdot N^n$ is the right time scale in the sense that

$$\mathcal{L}^{\pi_\theta}[\Theta_n(\mathbb{X}_n(t \cdot s_n))] \overset{n\to\infty}{\Longrightarrow} \mathcal{L}^\theta[Y_{2t}].$$

In order to stress the similarity to the stable case we might vaguely formulate our result as: "Given $\Theta_n(\mathbb{X}_n(t \cdot s_n)) = \rho$, the α-scaled finite system ${}^{f_\alpha}\mathbb{X}_n(t)$ relaxes fast to the 'equilibrium state' $\nu_\rho(\hat{\alpha})$ (defined in (8))." More precisely our Theorem II.4 says the following

$$\text{a)}\quad \mathcal{L}^{\pi_\theta}\left[\Theta_{f_\alpha(n)}(\mathbb{X}_n(t))\right] \overset{n\to\infty}{\Longrightarrow} \int_0^1 \mathbf{P}^\theta[Y_{2t} \in d\rho]\mathcal{L}^\rho[Y_{\hat{\alpha}}] = \mathcal{L}^\theta[Y_{2t+\hat{\alpha}}]$$

$$\text{b)}\quad \mathcal{L}^{\pi_\theta}\left[{}^{f_\alpha}\mathbb{X}_n(t \cdot s_n)\right] \overset{n\to\infty}{\Longrightarrow} \int_0^1 \mathbf{P}^\theta[Y_{2t} \in d\rho]\nu_\rho(\hat{\alpha}) = \int_0^1 \mathbf{P}^\theta[Y_{2t+\hat{\alpha}} \in d\rho]\pi_\rho$$

The finite systems scheme is not applicable to the strongly recurrent case of geometrical kernels $a_c,\ c > 1$. Here the local properties are no more completely determined by the macroscopic variable. In particular the process of block averages (going from large to small blocks) is no more Markov. As indicated above re-normalisation brings back the Markov property. Again we can recognise the chain. This will be done in Theorem II.5.

4.4 Branching Models

Once that we have understood the finite systems scheme and remembering the similarity of the results in Section 4 we will not be surprised to see the results for the branching models. In particular, the results by Cox and Greven (1990) concerning the finite systems scheme for branching random walks on $\mathbb{Z}^d,\ d \geq 3$, suggest that the scheme should apply to this setting.

Here the most interesting point seems to be the techniques employed for the proofs. We develop coupling techniques that allow a rather elegant handling of our branching processes. This shows a considerable progress over the methods used in Cox and Greven (1990). In particular there should be versions of this strategy in resampling systems of branching models with selection, mutation or even interaction between families (see Dawson, Greven and Vaillancourt (1995), Dawson and Greven (1995) and Cox, Dawson and Greven (1995)). However, the point of the proofs is deferred to the corresponding chapters. In particular, Section III.3 gives a general introduction to the coupling techniques used in this work.

Due to the similarity of the scenario, here we will content ourselves with a short description of the results.

Let $\Lambda_t^d = \mathbb{R}^d/(t\mathbb{Z}^d)$ be the d-dimensional torus of size $t > 0$. Consider a Brownian motion on Λ_t^d. As above we may now define the critically binary branching Brownian motion $(\eta_{t,s})_{s\geq 0}$ on Λ_t^d (shorthand BBM(Λ_t^d)). Also by the above high density limit procedure we obtain the super Brownian motion $(\zeta_{t,s})_{s\geq 0}$ on Λ_t^d (shorthand SBM(Λ_t^d)). These processes will be referred to as the *finite versions* of our brancing processes (allthough "bounded" might seem to be more appropriate in the continuum limit of the spatial structure). Either process will be denoted by $(\psi_{t,s})_{s\geq 0}$. We denote by $M_t(\rho)$, $\widetilde{M}_t(s)$, etc. the restrictions of $M(\rho)$, $\widetilde{M}(s)$ etc. to Λ_t^d. We define the rescaled finite systems $(\widetilde{\psi}^\alpha_{t,s})_{s\geq 0}$ as above by $\widetilde{\psi}^\alpha_{t,s}(B) = s^{-\alpha}\widetilde{\psi}^\alpha_{t,s}(s^{\alpha/2}B)$ and $\widetilde{\psi}_{t,s}(B) = \frac{8\pi}{\log s}\psi_{t,s}(B)$.

Consider first the stable case $d \geq 3$. Recall that (ν_ρ) are the equilibria for $(\psi_s)_{s\geq 0}$. The behaviour of the empirical population densities (block averages) is well known

$$\mathcal{L}^{M_t(\rho)}\left[t^{-d}\psi_{t,\sigma t^d}(\Lambda_t^d)\right] \overset{t\to\infty}{\Longrightarrow} Z_{\sigma/2}, \quad \sigma > 0,$$

where $(Z_s)_{s\geq 0}$ is Feller's diffusion (recall (13)). Theorem III.3 states

$$\mathcal{L}^{M_t(\rho)}\left[\psi_{t,\sigma t^d}\right] \overset{t\to\infty}{\Longrightarrow} \int_0^\infty \mathbf{P}^\rho[Z_{\sigma/2} \in d\rho']\nu_{\rho'}.$$

Continuing with $d = 2$ and letting $\beta(t) = t^2 \log t$ we see

$$\mathcal{L}^{\widetilde{M}_t(\sigma\cdot\beta(t))}\left[t^{-d}\widetilde{\psi}_{t,\sigma\beta(t)}(\Lambda_t^d)\right] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}^1[Z_{2\pi\sigma}].$$

(Here we made use of the basic scaling property for Feller's diffusion $\mathcal{L}^{\rho/\alpha}[\alpha Z_\beta] = \mathcal{L}^\rho[Z_{\alpha\beta}]$.)

Consequently, with the notation of Theorems III.1 and III.2 our Theorems III.4 and III.5 state

$$\mathcal{L}^{\widetilde{M}_t(\sigma\beta(t))}\left[(\widetilde{\psi}_{t,s}(s^{\alpha/2}x_k + [0,1[^2))_{k=1,\dots,n}\right] \overset{t\to\infty}{\Longrightarrow} \int_0^\infty \mathbf{P}^1[Z_{2\pi\sigma} \in d\rho]\mathcal{L}^\rho\left[(Z_1^{1-\alpha,k})_{k=1,\dots,n}\right],$$

for $\alpha \in [0,1]$ and

$$\mathcal{L}^{\widetilde{M}_t(\sigma\beta(t))}\left[(\widetilde{\psi}^\alpha_{t,s}([0,1[^2))_{\alpha\in I}\right] \underset{\text{fdd}}{\overset{t\to\infty}{\Longrightarrow}} \mathcal{L}^1\left[(Z_{2\pi\sigma+1-\alpha})_{\alpha\in I}\right].$$

Remark: It is known that $\beta(t) = t^2 \log t$ is the right time scale also to describe finite versions of other interacting particle systems in $\mathbb{R}^2$ and $\mathbb{Z}^2$. More generally, as in the case of the hierarchical group, the time scale has the form

volume of the box $\times$ recurrent potential maximised over the box.

It is again the blow-up technique employed here that emphasises this similarity.

Part II
Hierarchically Interacting Diffusions

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, i.e. 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 behaviour of large finite systems, indexed by finite subsets Ξ_n of Ξ, to that of the infinite one.

On the way we will establish relations between hitting times of random walks and their potentials.

1 Introduction and Main Results

1.1 Survey

In this part we analyse the pattern of cluster formation in systems of interacting diffusions and study the behaviour 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 been 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 (1993). 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 letting the interaction parameter approach its limit.

At the same time we are able to treat, in a likewise systematic way, the question of how the behaviours 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 (1995).

1.2 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 following system of stochastic differential equations (SSDE)

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

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 fulfill

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

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

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

with addition component wise modulo N ($N = 2,3\dots$ is some fixed parameter) and distance $\|\xi\| := \max\{k : \xi_k \neq 0\} \vee 0$. 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

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

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

$$r_k := a(\xi,\zeta)R_k \qquad \text{with} \qquad R_k := \#\{\xi\in\Xi : \|\xi\| = k\} = (N-1+1\!\!1_0(k))N^{k-1}. \tag{1.5}$$

This model has been 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 organised 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 $\xi = (\xi_1, \xi_2, \xi_3, \ldots)$ is the ξ_1th member of the ξ_2th family of the ξ_3th clan etc. 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 etc. 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. The flow spreads uniformly on the sites of degree k relatives ($k = 1, 2, \ldots$).

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

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

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, i.e.

$$\mathcal{L}^\mu[\mathbb{X}(t)] \overset{t\to\infty}{\Longrightarrow} \theta\delta_{\mathbf{1}} + (1-\theta)\delta_{\mathbf{0}}, \tag{1.7}$$

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

In the case a transient, 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 $\mu \in \mathcal{M}_\theta$

$$\mathcal{L}^\mu[\mathbb{X}(t)] \overset{t\to\infty}{\Longrightarrow} \nu_\theta. \tag{1.8}$$

(See Cox and Greven (1994a) Theorem 1 and 2)

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

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

1.3 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), 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 has been 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 $f : \mathbb{N}_0 \to \mathbb{N}_0, f(n) \uparrow \infty$ is called *space scale.* To keep time continuous we introduce the "inverse" of (s_n)

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

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

Let the shift operators $S_k : \Xi \longrightarrow \Xi$, $k = 0, 1, 2, \ldots$, 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

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

(ii) Block averages

For $n \in \mathbb{N}$ let the nth block average be defined by

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

The block averages are to be thought of as a macroscopic variable determining the behaviour 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

(i). Let $(Y_t)_{t\geq 0}$ be a standard Fisher-Wright diffusion on $[0,1]$, i.e. 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), Prop. 10.2.8). Hence $\lim_{t\to\infty} \mathbf{P}^\theta[Y_t = 1] = 1 - \lim_{t\to\infty} \mathbf{P}^\theta[Y_t = 0] = \theta$.

(ii). 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

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

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 e.g. 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

$$\log(k)\left[\log(r_k N^k) - \log(r_{k+1}N^{k+1})\right] \quad \text{is bounded.} \tag{1.15}$$

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

(iii). In the case a critical and for $\alpha \in [0,1]$ let the α-space-scale be a function $f_\alpha : \mathbb{N}_0 \to \mathbb{N}_0$ (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*

$$a)\quad \mathcal{L}^\mu\left[\Theta_{f_\alpha(n(t))}(\mathbb{X}(t))\right] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}^\theta[Y_{\hat\alpha}]$$

$$b)\quad \mathcal{L}^\mu\left[{}^{f_\alpha}(\mathbb{X}(t))\right] \overset{t\to\infty}{\Longrightarrow} \nu_\theta(\hat\alpha) := \int Q_{\hat\alpha}(\theta, d\rho)\pi_\rho$$

where $\mu \in \mathcal{M}_\theta$, $\hat\alpha := -\log\alpha$ *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. $\mathbb{X}$ converges towards $\nu(\hat\alpha)$ when rescaled with f_α and (s_n). The asymptotic behaviour 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

- $*$ small clusters when $\frac{f_\alpha(n)}{n} \overset{n\to\infty}{\longrightarrow} 0$ for $\alpha < 1$
- $*$ medium clusters when $\lim_{n\to\infty} \frac{f_\alpha(n)}{n} \in]0,1[$ for $\alpha \in]0,1[$
- $*$ large clusters when $\frac{f_\alpha(n)}{n} \overset{n\to\infty}{\longrightarrow} 1$ for $\alpha > 0$

For instance these above cases can occur if we choose

- $*$ $r_k = \vartheta k N^{-k}$ and $f_\alpha(n) = n^\alpha$
- $*$ $r_k = \vartheta N^{-k}$ and $f_\alpha(n) = \alpha n$
- $*$ $r_k = \vartheta k^{-\log k} N^{-k}$ and $f_\alpha(n) = n\left(1 + \frac{\log\alpha}{2\log n}\right)$

(ϑ some normalizing constants).

(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 behaviour 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 (1995).

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 to 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 neighbour interaction. Recently extensions have been made to Arratia's result which are concerned with stochastic partial differential equations models (Tribe (1993), Section 7) or more general interaction kernels in the voter model (Cox and Durrett (1994), Thm. 4). But these still rely on the linear structure. In order to circumvent this problem we use the idea of re-normalisation via block averages (recall (1.12)) and establish that the limiting density chain

$$(Z_m^{N,t}) = \mathrm{w} - \lim_{n\to\infty} \Theta_{n-m}(\mathbb{X}(ts_n)) \qquad \text{(with } m \text{ as time parameter)}$$

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

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

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

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

(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 $\kappa(N) \cdot c^n$ with $\kappa(N)(Nc)^{-1} \overset{N\to\infty}{\longrightarrow} 1$ we prefer therefore 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 non-negative martingale $(Z_m^{N,t})_{m\in\mathbb{Z}}$ such that*

$$\mathcal{L}^{\mu}\left[(\Theta_{n-m}(\mathbb{X}(ts_n)))_{m\in\mathbb{Z}}\right] \overset{n\to\infty}{\Longrightarrow} \mathcal{L}^{\theta}\left[(Z_m^{N,t})_{m\in\mathbb{Z}}\right]$$

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

$$\begin{array}{llcl} b) & \mathcal{L}^{\theta}[Z_m^{N,t}] & \overset{m\to\infty}{\Longrightarrow} & \theta\delta_1+(1-\theta)\delta_0 \\ & \mathcal{L}^{\theta}[Z_m^{N,t}] & \overset{m\to-\infty}{\Longrightarrow} & \delta_\theta \\ c) & (Z_m^t)_{m\in\mathbb{Z}} & := & \mathrm{w}-\lim_{N\to\infty}\left[(Z_m^{N,t})_{m\in\mathbb{Z}}\right] \text{ exists and is Markov.} \end{array}$$

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

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

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). $(Z_m^{N,t})$ 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 $N\to\infty$. (For more information on $(Z_m^{N,t})$ see Section 5.2, particularly (5.13) and (5.18).)

(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 $\mathcal{L}[X_\infty^\theta]$ one needs Baillon et al. (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 $f_\alpha(n)-n \overset{n\to\infty}{\longrightarrow} -\infty$ for $\alpha<1$, thus part b) of Theorem 2 would not hold.

Occupation times

We could take another approach to describe clustering phenomena, this is the investigation of occupation times. For $t\geq 0$ and $\xi\in\Xi$ we define the *occupation time at ξ up to time t* by

$$T_t(\xi)=\int_0^t x_\xi(s)\,ds. \tag{1.21}$$

We may investigate the asymptotic behaviour of $\frac{1}{t}T_t(\xi)$. Several questions have been dealt with in the literature such as: Does the limit exist? Is it non-degenerate? Does a central limit theorem hold? Can the space be properly re-scaled such that we obtain a non-trivial limiting field? These questions have been studied for the voter model in $\mathbb{Z}^d$ by Cox and Griffeath (1983), for critical branching Brownian motion on $\mathbb{R}^d$ by Cox and Griffeath (1985) and for super Brownian motion on $\mathbb{R}^d$ by Iscoe (1986). The answers to these questions are highly dimension dependent. In all cases it turns out that dimension $d = 2$ "is closer to" the cases $d \geq 3$ than to the case $d = 1$. We do not stress the point of occupation times in this work but only give a proposition that underlines this latter statement.

Proposition 1 *Let $\mu \in \mathcal{M}_\theta$ and $\theta \in [0, 1]$.*

(a). If the interaction kernel is strongly recurrent, i.e. a_c with $c > 1$, then

$$\mathrm{Var}^\mu[t^{-1}T_t(\xi)] \stackrel{t\to\infty}{\longrightarrow} 2\theta(1-\theta)(1 - 2^{-\log c/\log cN}). \tag{1.22}$$

(b). If the interaction kernel is critically recurrent then

$$\mathrm{Var}^\mu[t^{-1}T_t(\xi)] \stackrel{t\to\infty}{\longrightarrow} 0. \tag{1.23}$$

1.4 Finite Systems versus Infinite Systems

Since all computers known to the author so far (July 13, 2009) 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) and (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 $(\widetilde{Z}^{N,t}_m)$ 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

$$dx_{n,\xi}(t) = \left(\sum_{\zeta\in\Xi_n} a_n(\xi,\zeta)(x_{n,\zeta}(t) - x_{n,\xi}(t)) \right) dt + \sqrt{g(x_{n,\xi}(t))}\, dW_\xi(t) \quad (\xi \in \Xi_n), \tag{1.24}$$

where

$$a_n(\xi,\zeta) = \sum_{\substack{\zeta'\in\Xi \\ \zeta'\equiv\zeta(\mathrm{mod}\,\Xi_n)}} a(\xi,\zeta') \tag{1.25}$$

and

$$^f x_{n,\xi}(t) = x_\zeta(ts_n) \text{ where } \zeta = S^{-1}_{f(n)}\xi \tag{1.26}$$

$$\mathcal{L}^\mu[\mathbb{X}_n(0)] = \mu_n := \mu\Big|_{\Xi_n}. \tag{1.27}$$

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 heuristics 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

$$G(\xi, \zeta) = \sum_{m=0}^{\infty} a^{(m)}(\xi, \zeta). \tag{1.28}$$

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

$$V := \mathbf{E}^0 \left[\exp\left(-\frac{1}{2} \int_0^\infty \mathbb{I}_{\{X_s = 0\}} ds \right) \right] \tag{1.29}$$

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

Let the time scale be

$$s_n = \frac{G}{1-V} N^n. \tag{1.30}$$

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 holds

$$a) \quad \mathcal{L}^\mu\left[\Theta_n(\mathbb{X}_n(ts_n))\right] \overset{n\to\infty}{\Longrightarrow} \mathcal{L}^\theta[\mathrm{Y}_{2t}]$$

$$b) \quad \mathcal{L}^\mu\left[\mathbb{X}_n(ts_n)\right] \overset{n\to\infty}{\Longrightarrow} \int Q_{2t}(\theta, d\rho)\nu_\rho$$

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

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 holds

$$\begin{aligned} a)\quad & \mathcal{L}^\mu\left[\Theta_{f_\alpha(n)}(\mathbb{X}_n(t))\right] \overset{n\to\infty}{\Longrightarrow} \mathcal{L}^\theta[Y_{2t+\hat{\alpha}}] \\ b)\quad & \mathcal{L}^\mu\left[{}^{f_\alpha}\mathbb{X}_n(t\cdot s_n)\right] \overset{n\to\infty}{\Longrightarrow} \int Q_{2t}(\theta, d\rho)\nu_\rho(\hat{\alpha}) = \int Q_{2t+\hat{\alpha}}(\theta, d\rho)\pi_\rho \end{aligned}$$

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

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

$$\mathcal{L}^\mu\left[(\Theta_{n-m}(\mathbb{X}_n(ts_n)))_{m=0,1,\dots}\right] \overset{n\to\infty}{\Longrightarrow} \mathcal{L}^\theta\left[(\widetilde{Z}_m^{N,t})_{m=0,1,\dots}\right]$$

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

$$\begin{aligned} b)\quad & \mathcal{L}^\theta\left[\widetilde{Z}_m^{N,t}\right] \overset{m\to\infty}{\Longrightarrow} \theta\delta_1 + (1-\theta)\delta_0 \\ c)\quad & (\widetilde{Z}_m^t)_{m=0,1,\dots} := \mathrm{w} - \lim_{N\to\infty}\left[(\widetilde{Z}_m^{N,t})_{m=0,1,\dots}\right] \text{ exists and is Markov} \end{aligned}$$

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

$$\mathcal{L}\left[\widetilde{Z}_m^t | \widetilde{Z}_{m-1}^t = \rho\right] = \begin{cases} \mathcal{L}[Y_{2t}^\rho] & m = 0 \\ \mathcal{L}[X_\infty^\rho] & m > 0 \end{cases} \tag{1.31}$$

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 put 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.

1.5 Outline

The rest of the part is organised 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 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 1,3,4. Since Theorem 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 behaviour of hitting times of 0 for sequences of initial points which typically move away from 0. The key result is Proposition 2.7 in Subsection 2.4

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 $\Lambda_1 \subset \Lambda_2 \subset \dots$ and $\lim\limits_{n\to\infty} \Lambda_n = \Lambda$. E.g. think of $\Lambda = \mathbb{Z}^d$, $G_n = n\mathbb{Z}^d$ and $\Lambda_n = \big(]-\frac{n}{2}, \frac{n}{2}] \cap \mathbb{Z}\big)^d$. 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, i.e. $p_n(x,y) = \sum_{g\in G_n} p(x, y+g)$. By $(X(t))_{t\geq 0}$ resp. $(X_n(t))_{t\geq 0}$ denote the induced continuous time random walks, i.e. with transition probabilities

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

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

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

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

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 $G(x,y) = \sum\limits_{m=0}^{\infty} p^{(m)}(x,y)$ 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 with 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*

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

$$\sup_{x\in\Lambda_n} \left| |\Lambda_n| \cdot p_n^{([ts_n])}(0,x) - 1 \right| \overset{n\to\infty}{\longrightarrow} 0 \qquad \forall\, t > 0 \tag{2.11}$$

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

$$|\Lambda_n| \cdot \sup_{m \geq c_n |\Lambda_n| t} p^{(m)}(0,0) \overset{n\to\infty}{\longrightarrow} 0 \qquad \forall\, t > 0 \tag{2.12}$$

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

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

Here we used the notation $(c_n) \ll (a_n)$ for $\frac{c_n}{a_n} \overset{n\to\infty}{\longrightarrow} 0$.

2.2 Scaled Limits of Hitting Times

Assume $X(t)$ to be diffusive (either transient or recurrent) and let $(x_n)_{n\in\mathbb{N}}$ a sequence with $x_n \in \Lambda_n$, $n \in \mathbb{N}$ be such that

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

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

Proposition 2.2 (Diffusive Random Walk on Λ)

$$(i) \qquad \mathcal{L}^{x_n}\left[\frac{\tau}{s_n}\right] \stackrel{n\to\infty}{\Longrightarrow} (1-\alpha)\delta_0 + \alpha\delta_\infty$$

$$(ii) \quad \mathcal{L}^{x_n}\left[\frac{\tau_n}{s_n}\right] \stackrel{n\to\infty}{\Longrightarrow} (1-\alpha)\delta_0 + \alpha\cdot\mathcal{E}(1)$$

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

$$\begin{aligned} T_n(\lambda) &\stackrel{n\to\infty}{\longrightarrow} 1-\alpha \\ T'_n(\lambda) &\stackrel{n\to\infty}{\longrightarrow} 1-\alpha+\frac{\alpha}{1+\lambda}. \end{aligned}$$

By a simple first hitting time decomposition we obtain

$$T'_n(\lambda) = \frac{\sum_{m=0}^{\infty} p_n^{(m)}(0,x_n)e^{-\lambda m/s_n}}{\sum_{m=0}^{\infty} p_n^{(m)}(0,0)e^{-\lambda m/s_n}}. \tag{2.16}$$

We multiply by $\frac{1}{a_n}$ and split the dividend in three parts

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

The three sums are now estimated separately

(i)

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

since the integrand is bounded by $\frac{K}{|\Lambda_n|}e^{-\lambda t}$ and is $\sim \frac{1}{|\Lambda_n|}e^{-\lambda t}$ (by (2.11) and (2.12)).

(ii)

$$\begin{aligned} &\lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{\infty}\left[p^{(m)}(0,0) - p^{(m)}(0,x_n)\right]e^{-\lambda m/s_n} \\ &\stackrel{(2.14)}{=} \lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{c_n|\Lambda_n|}\left[p^{(m)}(0,0) - p^{(m)}(0,x_n)\right] \\ &\stackrel{(2.14)}{=} \lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{\infty}\left[p^{(m)}(0,0) - p^{(m)}(0,x_n)\right] = \alpha \end{aligned}$$

(iii)

$$\lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{\infty}p^{(m)}(0,0)e^{-\lambda m/s_n} \overset{(2.12)}{=} \lim_{n\to\infty}\frac{1}{a_n}\sum_{m=0}^{c_n|\Lambda_n|}p^{(m)}(0,0)e^{-\lambda m/s_n}$$
$$\overset{(2.13)}{=} 1$$

Putting the pieces together we obtain the convergence of the dividend to $\frac{1}{\lambda}-\alpha+1$. A similar expansion yields that the divisor converges to $\frac{1}{\lambda}+1$. 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. Then a_n can be chosen to be $\equiv G$ 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

$$(i) \quad \mathcal{L}^{x_n}\left[\frac{\tau}{s_n}\right] \overset{n\to\infty}{\Longrightarrow} \frac{\gamma}{G}\delta_0 + \left(1-\frac{\gamma}{G}\right)\delta_\infty$$

$$(ii) \quad \mathcal{L}^{x_n}\left[\frac{\tau_n}{s_n}\right] \overset{n\to\infty}{\Longrightarrow} \frac{\gamma}{G}\delta_0 + \left(1-\frac{\gamma}{G}\right)\mathcal{E}(1)$$

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

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

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 $|x_n| > b_n$*

$$\mathbf{P}^{x_n}(\tau_n/s_n > t) \overset{n\to\infty}{\longrightarrow} e^{-t}$$

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

$$\mathbf{P}^{x_n}(\tau_n/s_n > t) \overset{n\to\infty}{\longrightarrow} \alpha e^{-t}$$
$$\mathbf{P}^{x_n}(\tau/s_n > t) \overset{n\to\infty}{\longrightarrow} \alpha$$

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 Since $|\Lambda_n| = n^d$ we can choose $a_n = \frac{2}{\pi}\log n$ if $d = 2$ (see P12.3 of Spitzer (1964)). It remains to verify diffusiveness.

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

$$p^{(m)}(0,x) \leq K m^{-\frac{d}{2}} e^{-\frac{d|x|^2}{2m}} \tag{2.21}$$

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

$$m p^{(m)}(0,x) \leq K m^{1-\frac{d}{2}} \overset{m\to\infty}{\longrightarrow} 0$$

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

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

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

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

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

so

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

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.

Computation of the transition probabilities

Introduce

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

with addition componentwise modulo N and the scalar product

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

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

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

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

$$p(t;0,\xi) = (N-1) \sum_{k > \|\xi\|} N^{-k} e^{-t(1-f_k)} + (1\!\!1_{\{0\}}(\xi) - 1) N^{-\|\xi\|} e^{-t(1-f_{\|\xi\|})} \tag{2.26}$$

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

$$r_{n,k} = \begin{cases} r_k \left(1 - \sum\limits_{l=n+1}^{\infty} r_l\right)^{-1} & , \quad k \leq n \\ 0 & , \quad \text{else} \end{cases} \tag{2.27}$$

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

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

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

Note that (2.10) is always valid by symmetry.

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 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 let for $c > 1/N$ 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

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

$\mathcal{L}^\xi[T_n]$ coincide for all $\xi \in \Xi_n$. Hence by the Markov property $\mathcal{L}^\xi[T_n] = \mathcal{E}(\mu)$ for some $\mu \geq 0$ (recall $\mathcal{E}(\mu)$ 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 $t \to 0$

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

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

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

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

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

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

Case $X(t)$ critical

Recall that a recurrent random walk on Ξ is called critical if

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

This implies

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

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)

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

Note that in particular

$$\frac{A(n+1)}{A(n)} \overset{n\to\infty}{\longrightarrow} 1. \tag{2.42}$$

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

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

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 in 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 obtains a maximum of value $\leq \frac{2}{\varepsilon \log n}$ at $k_0 = n + \frac{\log(2\varepsilon)+\log\log n}{\log N}$ 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)

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

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

□

Key result on hitting times

Up to now we have proved the following

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

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

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

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

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

$$\mathbf{P}^{\xi_n}\left[\frac{\tau}{s_n} > t\right] \overset{n\to\infty}{\longrightarrow} \alpha \tag{2.47}$$

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

□

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

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

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

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

3 Coalescing Random Walks

We introduce the notion 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, i.e. one of the two particles dies and the other goes on moving. Call this new process $\widetilde{\eta}(t)$ the *system of instantaneously coalescing random walks.* Finally change the coalescence mechanism by not letting coalescence occur instantaneously but at a constant rate $b > 0$. This is a pair of particles coalesces after the particles have spent together an exponential waiting time with mean $\frac{1}{b}$. Call this new process a *system of delayed coalescing random walks* (with delay $\frac{1}{b}$) and denote it by $\eta(t)$. We are interested in $\eta(t)$ because of the mentioned duality relation. Since $\widetilde{\eta}(t)$ is easier to handle we first investigate this and then compare $\widetilde{\eta}(t)$ with $\eta(t)$. By $\overline{X}_n$, $\eta_n(t)$, $\widetilde{\eta}_n(t)$ etc we denote the corresponding objects on Ξ_n.

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

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

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

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

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

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

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

$\widetilde{\eta}(t)$ preserves $\widetilde{\Phi}_m := \widetilde{\Phi} \cap \Phi_m$ and on this has generator $\mathcal{H}_m$ defined for $f \in C_\mathrm{b}(\widetilde{\Phi}_m)$ by

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

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 $\xi_{n,1}, \ldots, \xi_{n,m}$, i.e. in

$$\varphi_n := 1\!\!\mathrm{I}_{\xi_{n,1}} + \ldots + 1\!\!\mathrm{I}_{\xi_{n,m}} \tag{3.5}$$

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

$$\frac{A(\xi_{n,i}\,,\,\xi_{n,j})}{A(n)} \stackrel{n\to\infty}{\longrightarrow} \alpha \qquad \forall\, i \neq j. \tag{3.6}$$

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

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 $q_t(m;m) = e^{-\binom{m}{2}t}$ and recall $\hat{\alpha} = -\log\alpha$.

Proposition 3.2 (Scaling Limit, Infinite Case)

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

We introduce the following notations

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

In view of Corollary 2.8 is 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 the "meeting probability" is given by our quantity α. Namely we show

Lemma 3.3

$$\mathbf{P}^{\varphi_n}[\overline{\tau} \leq s_n] \stackrel{n\to\infty}{\longrightarrow} 1 - \alpha^{\binom{m}{2}}. \tag{3.10}$$

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))

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

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

$$\mathbf{P}^{\xi_n}\left[\tau > t(\gamma, n)\right] \stackrel{n\to\infty}{\longrightarrow} \frac{\alpha}{\gamma} \tag{3.12}$$

where we put $t(\gamma, n) := \frac{1}{r_{f_\gamma(n)}}$. 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

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

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

$$\mathbf{P}^{\xi_n}[X(t(\gamma, n)) \in \Xi(\gamma, n)] \stackrel{n\to\infty}{\longrightarrow} 1 \tag{3.13}$$

since

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

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

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

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

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

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

$$= \int_\alpha^1 \sum_{\xi\in\Xi} \mathbf{P}^{\varphi_n}[\overline{\tau} = \tau(1,2) \in dt(n,\gamma),\ X(2, t(n,\gamma)) = \xi] \cdot \mathbf{P}[X(3, t(n,\gamma)) - \xi \notin \Xi(n,\gamma)]$$

by symmetry and (3.13)

$$= \int_\alpha^1 \sum_{\xi\in\Xi} \mathbf{P}^{\varphi_n}[\overline{\tau} = \tau(1,2) \in dt(n,\gamma),\ X(2, t(n,\gamma)) = \xi]$$

$$\cdot \mathbf{P}[X(3, t(n,\gamma)) \notin \Xi(n,\gamma)] + \varepsilon(n)$$

$= \varepsilon(n)$ by dominated convergence.

Now we put the pieces together

$$\mathbf{P}^{\varphi_n}\left[\tau(i,j) \le \frac{1}{r_n}\right] = \mathbf{P}^{\varphi_n}\left[\overline{\tau} = \tau(i,j) \le \frac{1}{r_n}\right] + \sum_{\{k,l\}\neq\{i,j\}} \int_{t(\alpha,n)}^{t(1,n)} \mathbf{P}^{\varphi_n}[\overline{\tau} = \tau(k,l) \in dt, \tau(i,j) \le t(1,n)] + \varepsilon(n) \tag{3.17}$$

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

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

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

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

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

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

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

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

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)

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

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 $\frac{s_n'}{s_n} \overset{n\to\infty}{\longrightarrow} 0$ slowly enough that

$$\mathbf{P}^{\varphi_n}[X(i,u) \in \Xi_n\ \forall\, u \le s_n', \forall\, i \text{ and } X(i,s_n') - X(j,s_n') \in \Xi(1,n)\,\forall\, i \neq j] \overset{n\to\infty}{\longrightarrow} 1 \tag{3.21}$$

and

$$\mathbf{P}^{\varphi_n}[\#\widetilde{\eta}(ts_n) = k] \overset{n\to\infty}{\longrightarrow} q_{\hat{\alpha}}(m;k).$$

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

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

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 $\overline{\tau}_n$ are the finite objects corresponding to those defined in (3.9))

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

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

$$\mathbf{P}^{\varphi_n}[\overline{\tau}_n \le ts_n] \overset{n\to\infty}{\longrightarrow} 1 - e^{-2t\binom{m}{2}}. \tag{3.24}$$

Now the induction argument cited above yields

$$\mathbf{P}^{\varphi_n}[\#\widetilde{\eta_n}(ts_n) = k] \overset{n\to\infty}{\longrightarrow} q_{2t}(m;k). \tag{3.25}$$

□

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

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

$$\mathbf{P}\left[\widetilde{\eta_n}^{\varphi_n^*}(ts_n) = \eta_n^{\varphi_n}(ts_n)\right] \overset{n\to\infty}{\longrightarrow} 1, \qquad t > 0. \tag{3.27}$$

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

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

the time points of coalescence and note that

$$T_{i+1}^n - T_i^n \overset{n\to\infty}{\longrightarrow} \infty \qquad \mathbf{P}\text{-a.s.}$$

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 Proposition 3.2 and 3.4 and Lemma 3.5 we have proved

Proposition 3.6 (Scaling Limits)

$$\begin{aligned} \mathcal{L}^{\varphi_{n(t)}}[\#\eta(t)] &\overset{t\to\infty}{\Longrightarrow} \mathcal{L}^{m^*}[D_{\hat{\alpha}}] \\ \mathcal{L}^{\varphi_n}[\#\eta_n(ts_n)] &\overset{n\to\infty}{\Longrightarrow} \mathcal{L}^{m^*}[D_{2t+\hat{\alpha}}] \quad \textit{for } t>0 \textit{ fixed.} \end{aligned}$$

□

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 did depend 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 $t_n \uparrow \infty$, $t_n \ll \left(\sum_{k>n} r_k\right)^{-1}$, then (by (2.30)) $t_n \ll N^{-n}$ and

$$\mathbf{P}\left[X_n(t) = X(t)\ \forall\ t \le t_n\right] \overset{n\to\infty}{\longrightarrow} 1. \tag{3.28}$$

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

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

Since

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

and by Proposition 2.7 we get

$$\mathcal{L}\left[\frac{\tau_n^{(i+1)} - \tau_n^{(i)}}{G|\Xi_n|}\right] \overset{n\to\infty}{\Longrightarrow} \mathcal{E}(1). \tag{3.31}$$

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

$$\mathcal{L}\left[\max\left\{k : \frac{\tau_n^{(k)}}{G|\Xi_n|} \le t\right\}\right] \overset{n\to\infty}{\Longrightarrow} \mathcal{L}[B(t)]. \tag{3.32}$$

Recall

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

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

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

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

Proposition 3.7

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

□

4 Proof of Theorem 1,3 and 4

4.1 Proof of Theorem 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 in the product measure π_θ and where $g(x) = bx(1-x),\ b > 0$.

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 superscript where necessary. Let now $\eta(t)$ be a system of coalescing random walks with delay $\frac{1}{b}$ and let

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

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

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

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 $\varphi = k_1 1\!\!1_{\xi_1} + \ldots + k_{m^*} 1\!\!1_{\xi_{m^*}} \in \Phi,\ m^* \in \mathbb{N},\ k_1,\ldots,k_{m^*} \in \mathbb{N},\ \xi_i \neq \xi_j$, i.e. 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

$$\mathbf{E}^{\pi_\theta}\left[\left({}^{f_\alpha}\mathbb{X}^b(t)\right)^{\varphi}\right] = \mathbf{E}^{\pi_\theta}\left[\left(\mathbb{X}^b(t)\right)^{\varphi_{n(t)}}\right] \overset{t\to\infty}{\longrightarrow} \mathbf{E}^{\theta}\left[(Y_{\hat\alpha})^{m^*}\right] \tag{4.2}$$

$$\mathbf{E}^{\pi_\theta}\left[\left({}^{f_\alpha}\mathbb{X}^b_n(n)\right)^{\varphi}\right] = \mathbf{E}^{\pi_\theta}\left[\left(\mathbb{X}^b_n(ts_n)\right)^{\varphi_n}\right] \overset{n\to\infty}{\longrightarrow} \mathbf{E}^{\theta}\left[(Y_{2t+\hat\alpha})^{m^*}\right] \tag{4.3}$$

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

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

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 the same.

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, i.e. we show that two versions $\mathbb{X}^1$ and $\mathbb{X}^2$ of our interacting system with ergodic initial laws μ and ν with same intensity θ can be defined on one probability space such that $\mathbf{E}[|x_0^1(t) - x_0^2(t)|] \overset{t\to\infty}{\longrightarrow} 0$. Define the four-valued process $(\mathbb{X}^1, \mathbb{X}^2, \mathbb{X}_n^1, \mathbb{X}_n^2)$ as the solution of

$$dx_\xi^i(t) = \sum_{z\in\Xi} a(\xi,\zeta)(x_\zeta^i(t) - x_\xi^i(t))dt + \sqrt{bx_\xi^i(1-x_\xi^i)}\, dW_\xi(t), \quad i = 1,2 \tag{4.4}$$

$$dx_{n,\xi}^i(t) = \sum_{z\in\Xi_n} a(\xi,\zeta)(x_{n,\zeta}^i(t) - x_{n,\xi}^i(t))dt + \sqrt{bx_{n,\xi}^i(1-x_{n,\xi}^i)}\, dW_\xi(t), \quad i = 1,2 \tag{4.5}$$

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

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

and

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

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

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

$$\mathbf{E}[|\Delta_0(t)|] \overset{t\to\infty}{\longrightarrow} 0 \tag{4.6}$$

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 this

Lemma 4.2 (Successful coupling, Finite systems)
Under the same conditions as in Lemma 4.1

$$\mathbf{E}[|\Delta_{n,0}(ts_n)|] \overset{n\to\infty}{\longrightarrow} 0 \tag{4.7}$$

Proof Since the infinite systems can be coupled successfully we have to show that the finite and the infinite system do not diverge for sufficiently large time and that finite systems stay close once that they got 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

$$\sup_{n\geq m} \mathbf{E}[|\Delta_\xi^n(t_m)|] \overset{m\to\infty}{\longrightarrow} 0. \tag{4.8}$$

To see this we may proceed as Yamada and Watanabe (1971). We approximate the $|\cdot|$-function by functions $f_n(x) = \sqrt{\frac{1}{n} + x^2}$ 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

$$\begin{aligned} d\mathbf{E}|\Delta^n_\xi(t)| &\leq \mathbf{E}\left[\sum_{\zeta\in\Xi_n} a(\xi,\zeta)(|\Delta^n_\zeta(t)| - |\Delta^n_\xi(t)|)\right] dt \\ &+ \mathbf{E}\left[\sum_{\zeta\notin\Xi_n} a(\xi,\zeta)(|x_\zeta(t)| + |x_{\xi_{|\Xi_n}}(t)|)\right] dt \end{aligned} \tag{4.10}$$

The first term vanishes by translation invariance ($\Xi_n \trianglelefteq \Xi$ subgroup!) 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, i.e. $\mathbf{E}[|\Delta_0(t_n)|] = \varepsilon(n)$, and so are the finite systems. Hence it is enough to show

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

This is however true since as above

$$d\mathbf{E}[|\Delta_{n,\xi}(t)|] \leq \mathbf{E}\left[\sum_{\zeta\in\Xi_n} a(\xi,\zeta)(|\Delta_{n,\zeta}(t)| - |\Delta_{n,\xi}(t)|)\right] dt = 0 \tag{4.12}$$

□

Generalization to Admissable $g(x)$

Finally we generalize the diffusion coefficient. Fix an admissable 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 behaviour 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

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

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

$$\mathbf{E}^\mu\left[(\mathbb{X}^{g^\varepsilon}(t))^\varphi\right] \leq \mathbf{E}^\mu\left[(\mathbb{X}^{g}(t))^\varphi\right] \leq \mathbf{E}^\mu\left[(\mathbb{X}^{bf}(t))^\varphi\right] \qquad \forall\, t \geq 0, \tag{4.13}$$

which is due to Cox, Fleischmann and Greven (1995), Theorem 1.

We introduce the linear map

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

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

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

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

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

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

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

Recall the definition of p_φ from (3.34) and note that

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

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

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

5 Proof of Theorem 2 and 5

We only consider the case $g(x) = x(1-x)$ and $\mathcal{L}[\mathbb{X}(0)]$ 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 part a) and b) of Theorem 2 and 5. Then we obtain c) via a duality to the discrete time nonlinear death process of Definiton 3.1.

5.1 Limit Process of Scaled Random Walks

In this subsection we "extend Ξ resp. Ξ_n to the left", i.e. 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*

$$\Gamma := \{\delta = (\delta_k)_{k\in\mathbb{Z}} : \delta_k \in \{0,\dots,N-1\}\,,\ \|\delta\| < \infty\} \tag{5.1}$$
$$\Gamma_{-n} := \{\delta \in \Gamma : \delta_k = 0\ \forall\, k \le -n\} \tag{5.2}$$

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

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

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

Further let μ *resp.* μ' *be the Haar measures on* Γ *resp.* Γ' *normed to* $\mu(\Gamma') = \mu'(\Gamma') = 1$ *(sic!), i.e. 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 occuring differences.

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

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

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

The same way we obtain the system of instantaneously coalescing random walks $\overline{\beta}_n$ on Γ

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

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 $(\widehat{G}_n f)\Big|_{\Gamma_{-n}} = G_n(f\Big|_{\Gamma_{-n}})$. 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 metrics 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 $C(\Gamma)$.

Definition 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* $\overline{\beta}(\delta_1,\dots,\delta_m;t)$ *we denote the corresponding system of of instantaneously coalescing random walks started in* $(\delta_1,\dots,\delta_m)$.

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 $\mathcal{R}(\lambda - \widehat{G}_n) = \widehat{C}(\Gamma_{-n})$. So the range of $\lambda - G$ is dense, $\mathcal{R}(\lambda - G) = \widehat{C}(\Gamma)$, and hence G is recognized as a Markov generator. (For a treatment of this point see Liggett (1985), Chapter I). □

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

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

Now it is immediate that

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

Lemma 5.3

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

Proof Let

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

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

Now by (5.8) and right continuity of paths

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

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

$$\overline{\beta}_1 \geq \overline{\beta}_2 \geq \ldots \overline{\beta} \tag{5.12}$$

and since τ has no atoms a simple induction argument yields the conclusion. □

5.2 Proof of Theorems 2 and 5, Part a)

By compactness of the state space it suffices to show convergence of moments

$$\mathbf{E} \prod_{m\in\mathbb{Z}} \left(\Theta_{n-m}(\mathbb{X}(ts_n))\right)^{\psi_m} \stackrel{n\to\infty}{\longrightarrow} M_\psi \tag{5.13}$$

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

Thus let

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

and denote

$$\Gamma(\psi) = \left\{(\overline{\delta} = (\delta_1, \ldots, \delta_r) \in \Gamma^r : \|\delta_i\| \leq l_i\right\} \tag{5.15}$$

$$\Gamma_{-n}(\psi) = \Gamma(\psi) \cap (\Gamma_{-n})^r.$$

Then

$$\mathbf{E}\left[(\Theta_{n-\cdot}(\mathbb{X}(ts_n)))^{\psi}\right] = \mathbf{E}\left[\prod_{j=1}^{r}\Theta_{n-l_j}(\mathbb{X}(ts_n))\right] \tag{5.16}$$

$$= \left(\prod_{j=1}^{r}\#\Xi_{n-l_j}\right)^{-1}\mathbf{E}\left[\sum_{\|\xi_1\|\leq n-l_1}\cdots\sum_{\|\xi_r\|\leq n-l_r} x_{\xi_1}(t(Nc)^{n+1})\cdots x_{\xi_r}(t(Nc)^{n+1})\right]$$

By the duality lemma and the comparison lemma this equals

$$= \left(\prod_{j=1}^{r}\#\Xi_{n-l_j}\right)^{-1}\mathbf{E}\left[\sum_{\|\xi_1\|\leq n-l_1}\cdots\sum_{\|\xi_r\|\leq n-l_r}\theta^{\#\widetilde{\eta}_n^{\{\xi_1,\dots,\xi_r\}^*}(t(Nc)^{1+n})}\right] + \varepsilon(n) \tag{5.17}$$

$$= \int \mathbf{E}^{\overline{\delta}^*}\left[\theta^{\#\overline{\beta}_n(t)}\right]\mu^r\left(d\overline{\delta}|\Gamma_{-n}(\psi)\right) + \varepsilon(n).$$

By Lemma 5.3 this tends to

$$M_{\psi} := \int \mathbf{E}^{\overline{\delta}^*}\left[\theta^{\#\overline{\beta}(t)}\right]\mu^r\left(d\overline{\delta}|\Gamma(\psi)\right). \qquad \square \tag{5.18}$$

5.3 Proof Theorems 2 and 5, Part b)

We do the proof by an explicit calculation using Laplace transforms of the first hitting times τ.

It suffices to show (recall τ from (5.10) and note that here d plays the role of m in Theorem 2 and 5)

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

since then (recall M_ψ from (5.18))

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

A straightforward computation using (2.16) and abbreviating $v = \frac{Nc}{Nc-1} + \frac{1}{N-1}$ yields

$$\mathbf{E}^d e^{-\lambda\tau} = \lim_{n\to\infty}\mathbf{E}^d e^{-\lambda\tau_n} = \frac{\displaystyle\sum_{m=d-1}^{\infty}\frac{N^{-m}}{\vartheta v(Nc)^{-m}+\lambda} - \frac{N}{N-1}\frac{N^{1-d}}{\vartheta v(Nc)^{1-d}+\lambda}}{\displaystyle\sum_{m=-\infty}^{\infty}\frac{N^{-m}}{\vartheta v(Nc)^{-m}+\lambda}} \tag{5.20}$$

whereas

$$\mathbf{E}^d e^{-\lambda\tau'} = \lim_{n\to\infty}\mathbf{E}^d e^{-\lambda\tau'_n} \tag{5.21}$$

$$= \frac{\displaystyle\sum_{m=d}^{0}\frac{N^{1-m}}{v(Nc-1)(Nc)^{-m}-1+\lambda} - \frac{N}{N-1}\frac{N^{1-d}}{v(Nc-1)(Nc)^{1-d}-1+\lambda} + \frac{N}{N-1}\frac{1}{\lambda}}{\displaystyle\sum_{m=-\infty}^{0}\frac{N^{1-m}}{v(Nc-1)(Nc)^{-m}-1+\lambda} + \frac{N}{N-1}\frac{1}{\lambda}}.$$

Now (5.19) follows from

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

and

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

5.4 Proof Theorems 2 and 5, Part c)

We give a qualitative description of a system of coalescing random walks in the limit $N\to\infty$. (A verbal description is given below (5.33).) Then we construct a process dual to the Fisher-Wright diffusion with immigration (Definition 5.4). This will serve to conclude the proof via a moment calculation.

We let $N\to\infty$ and indicate quantities with a superscript N. Observe

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

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

Thus

$$\mathbf{E}^d e^{-\lambda\tau^N(Nc)^{-a}}\overset{N\to\infty}{\longrightarrow}\begin{cases}1-c^{d-a}+\dfrac{(c-1)c^{d-a-1}}{1+\lambda/c} & if\quad d-a\le 0\\ 0 & if\quad d-a>0\end{cases} \tag{5.26}$$

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

$$\mathbf{E}^d e^{-\lambda\tau'^N}\overset{N\to\infty}{\longrightarrow}1-c^d+\frac{c^d}{1+\frac{\lambda}{c-1}}. \tag{5.27}$$

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

$$\mathcal{L}^d[\tau^N(Nc)^{-a}]\overset{N\to\infty}{\Longrightarrow}\begin{cases}\left(1-c^{d-a}\right)\delta_0+(c-1)c^{d-a-1}\mathcal{E}\left(\frac{1}{c}\right)+c^{d-a-1}\delta_\infty & \text{if}\quad d\le a\\ \delta_\infty & \text{if}\quad d>a\end{cases} \tag{5.28}$$

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

$$\mathcal{L}^d[\tau'^N]\overset{N\to\infty}{\Longrightarrow}\left(1-c^d\right)\delta_0+c^d\mathcal{E}\left(\frac{1}{c-1}\right). \tag{5.29}$$

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

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

As in (3.14) we obtain

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

Thus

$$\lim_{N\to\infty} \mathbf{P}^{\delta}[\|\gamma(t(Nc)^n\| = 1+n] = \qquad (5.32)$$
$$1 - \lim_{N\to\infty} \mathbf{P}^{\delta}[\|\gamma(t(Nc)^n\| =\leq n] = 1 - e^{-t} \quad \text{if}\|\delta\| \leq n.$$

By (5.28)

$$\mathbf{P}^d\left[\sigma_a^N \leq \tau^N \leq t(Nc)^a\right] \overset{N\to\infty}{\longrightarrow} 0$$

and thus

$$\mathbf{P}^d\left[\tau^N \leq \sigma^N\right] \overset{N\to\infty}{\longrightarrow} \frac{c-1}{c}.$$

Hence we get

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

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 $\psi = 1\!\!1_{d_1} + \ldots + 1\!\!1_{d_r}$ as above let

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

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

$$L_d^N(t) \;=\; \#\{\delta \in \overline{\beta}^N(t) : \|\delta\| \leq d\} \qquad (5.34)$$
$$U_d^N(t) \;=\; \#\{\delta \in \overline{\beta}^N(t) : \|\delta\| > d\}. \qquad (5.35)$$

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 $2(c-1)(Nc)^{-d_1}$. Independent 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 behaviour of this will be modelled by

Definition 5.4 (Death-Escape Process)
Let (A_t, B_t) be the $\mathbb{N}\times\mathbb{N}$-valued Markov process with generator

$$\mathcal{G}\left((a_1,b_1),(a_2,b_2)\right) = \begin{cases} 2(c-1)\binom{a_1}{2} & \text{if} \quad a_2 = a_1 - 1 \;,\; b_2 = b_1 \\ a_1 & \text{if} \quad a_2 = a_1 - 1 \;,\; b_2 = b_1 + 1 \\ -a_1 - 2(c-1)\binom{a_1}{2} & \text{if} \quad a_2 = a_1 \qquad ,\; b_2 = b_1 \end{cases} \qquad (5.36)$$

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 towards θ in (1.17). One easily checks the following duality relation

Lemma 5.5 (Duality)

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

□

Let

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

Then

$$\mathcal{L}^{\psi}\left[L_{d_1}^N(t(Nc)^{d_1}), U_{d_0}^N(t(Nc)^{d_1})\right] \overset{N\to\infty}{\Longrightarrow} \mathcal{L}^{\psi}\left[(X_t, Y_t)|X_0 = \psi_{d_1}, Y_0 = \psi_{d_1}^+\right].$$

Thus

$$\mathcal{L}^{\psi}\left[\#\overline{\beta}^N(t(Nc)^{d_1})\right] \overset{N\to\infty}{\Longrightarrow} \mathcal{L}^{\psi}\left[\psi_{d_1}^+ + G_\infty(\psi_{d_1})\right].$$

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

$$\mathcal{L}^{\psi}\left[\#\overline{\beta}^N\left(Nc^{-0.5}\right)\right] \overset{N\to\infty}{\Longrightarrow} \mathcal{L}^{\psi}\left[\psi_{-1}^+ + G_\infty(\psi_{-1} + G_\infty(\psi_{-2} + \ldots + G_\infty(\psi_{d_1})\cdots).\right]$$

Finally we get for the infinite system

$$\mathcal{L}^{\psi}\left[\#\overline{\beta}^N(t)\right] \overset{N\to\infty}{\Longrightarrow} \mathcal{L}^{\psi}\left[\psi_0^+ + G_t(\psi_0 + G_\infty(\psi_{-1} + G_\infty(\psi_{-2} + \ldots + G_\infty(\psi_{d_1})\cdots)\right]. \quad (5.38)$$

In the last step the finite system differs from the infinite one since in the former is $\sigma_0 \equiv \infty$ 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

$$\mathcal{L}^{\psi}\left[\#\overline{\beta}'^N(t)\right] \overset{N\to\infty}{\Longrightarrow} \mathcal{L}^{\psi}\left[\psi_0^+ + G_t'(\psi_0 + G_\infty(\psi_{-1} + G_\infty(\psi_{-2} + \ldots + G_\infty(\psi_{d_1})\cdots)\right]. \quad (5.39)$$

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

$$q_t(\psi, m) = \mathbf{P}^{\psi}\left[\psi_0^+ + G_t(\psi_0 + G_\infty(\psi_{-1} + G_\infty(\psi_{-2} + \ldots + G_\infty(\psi_{d_1})\cdots) = m\right] \quad (5.40)$$

in (5.38) and (5.39) respectively and observe

$$\left(\mu^N\right)^r(\Delta(\psi)|\Gamma(\psi)) \overset{N\to\infty}{\longrightarrow} 1. \quad (5.41)$$

Hence

$$M_\psi^N \overset{N\to\infty}{\longrightarrow} \sum_m q_t(\psi, m)\theta^m \quad (5.42)$$

and

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

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 (1.31) 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 Theorem 2 and 5, part c).

□

6 The Behaviour of the Occupation Times

We investigate the asymptotic behaviour of the re-normed occupation time $t^{-1}T_t(\xi)$ defined by $T_t(\xi) := \int_0^t x_\xi(s)\,ds$, where $\mathbb{X}(s) = (x_\xi(s),\, \xi \in \Xi)$ is a system of linearly interacting *Fisher-Wright* diffusions on the hierarchical group Ξ.

Recall that $a_c,\ c > \frac{1}{N}$ with $r_k = \vartheta_c \cdot (Nc)^{-k}$ ($\vartheta_c = \frac{Nc-1}{Nc}$) are the geometrical kernels. a_c is strongly recurrent if $c > 1$ and critically recurrent if $c = 1$. Recall also that $\mathcal{M}_\theta$ is the class of ergodic measures with intensity θ introduced in (1.6).

Proposition 6.1 *Let $\mu \in \mathcal{M}_\theta$, $\theta \in [0,1]$.*

(a). If the interaction kernel is strongly recurrent, i.e. a_c with $c > 1$, then

$$\mathrm{Var}^\mu[t^{-1}T_t(\xi)] \overset{t\to\infty}{\longrightarrow} 2\theta(1-\theta)(1 - 2^{-\log c/\log cN}). \tag{6.1}$$

(b). If the interaction kernel is critically recurrent then

$$\mathrm{Var}^\mu[t^{-1}T_t(\xi)] \overset{t\to\infty}{\longrightarrow} 0. \tag{6.2}$$

We prepare for the proof with a couple of lemmas.

The case a strongly recurrent

First consider the case $a = a_c,\ c > 1$ is strongly recurrent.

From (2.26) we obtain the transition probabilities for the continuous time random walk generated by $a_c(\cdot,\cdot)$

$$\begin{aligned} p(t,0,\xi) &= (N-1) \sum_{k=\|\xi\|+1}^{\infty} N^{-k} \exp\left\{-tv(Nc)^{-k}\right\} \\ &\quad + (\mathbb{1}_{\{0\}}(\xi) - 1) N^{-\|\xi\|} \exp\left\{-tv(Nc)^{-\|\xi\|}\right\}, \end{aligned} \tag{6.3}$$

where

$$v := \frac{Nc}{Nc-1} + \frac{1}{N-1}.$$

We introduce the following notation

$$\gamma := \sum_{k=-\infty}^{\infty} N^{-k} \exp\left\{-v(Nc)^{-k}\right\} \tag{6.4}$$

$$\alpha := \frac{\log c}{\log Nc}. \tag{6.5}$$

Lemma 6.2 $\qquad p(t,0,0) \sim \gamma t^{\alpha-1}, \qquad t \to \infty.$

Proof First let $T \in \mathbb{N}$. Then we have

$$N^T p((Nc)^T, 0, 0) = (N-1) \sum_{k=1-T}^{\infty} N^{-k} \exp\left\{-(Nc)^{-k}\right\} \overset{T\to\infty}{\longrightarrow} \gamma. \tag{6.6}$$

By a simple monotonicity argument (6.6) holds for $T \in [0,\infty[$. Put $t = (Nc)^T$. Then $N^T = t^{1-\alpha}$ and the proof is complete. $\square$

Corollary 6.3 *Let* $0 < \beta < (\log Nc)^{-1}$. *Then*

$$\sum_{\|\xi\| \le \beta \log t} p(t,0,\xi) \stackrel{t\to\infty}{\longrightarrow} 0.$$

Proof

$$\begin{aligned}
\limsup_{t\to\infty} \sum_{\|\xi\| \le \beta \log t} p(t,0,\xi) &\le \limsup_{t\to\infty} \sum_{\|\xi\| \le \beta \log t} p(t,0,0) \\
&= \limsup_{t\to\infty} N^{\beta \log t} \cdot \gamma t^{\alpha-1} \\
&= \limsup_{t\to\infty} \gamma \cdot t^{\beta \log N + \alpha - 1} = 0,
\end{aligned}$$

since by assumption $\beta \log N + \alpha - 1 < 0$. □

Denote by $\mathbf{P}^\xi[\cdot]$ the probability associated with the random walk $\eta(t)$ generated by $a(\cdot,\cdot) = a_c(\cdot,\cdot)$ and with start in $\xi \in \Xi$. Let also

$$\mathbf{P}_0[\cdot] = \frac{1}{1 - a(0,0)} \sum_{\xi \neq 0} a(0,\xi) \mathbf{P}^\xi[\cdot] \tag{6.7}$$

the distribution of $\eta(t)$ after first leaving the origin. Let

$$\tau_0 = \inf\{t \ge 0 \,:\, \eta(t) = 0\}$$

the first hitting time the origin. In order to study the asymptotics of $\mathbf{P}_0[\tau_0 > t]$ as $t \to \infty$ we introduce the Laplace transforms

$$L(\lambda) := \int_0^\infty e^{-\lambda t} p(t,0,0)\, dt \tag{6.8}$$

$$L_0(\lambda) := \int_0^\infty e^{-\lambda t} \mathbf{P}_0[\eta(t) = 0]\, dt. \tag{6.9}$$

Lemma 6.4 *Denote by* $\Gamma(x)$ *the ordinary gamma function. Then*

$$L(\lambda) \sim L_0(\lambda) \sim \frac{\gamma}{\alpha} \Gamma(1+\alpha) \lambda^{-\alpha}, \qquad \lambda \to 0.$$

Proof

$$\int_0^t p(s,0,0)\, ds \sim \frac{\gamma}{\alpha} t^\alpha, \qquad t \to \infty,$$

by Lemma 6.2. By a Tauberian theorem (see e.g. Feller (1966), XIII.5, Thm. 1) we get

$$L(\lambda) \sim \frac{\gamma}{\alpha} \Gamma(1+\alpha) \lambda^{-\alpha}, \qquad \lambda \to 0. \tag{6.10}$$

The proof of $L(\lambda) \sim L_0(\lambda)$ may be done as in Fleischmann and Greven (1994), proof of Lemma 2.33. We omit the details. □

Lemma 6.5

$$\mathbf{P}_0[\tau_0 > t] \sim \frac{1-\alpha}{\frac{\gamma}{\alpha}\Gamma(1+\alpha)\Gamma(2-\alpha)}\, t^{-\alpha}, \qquad t \to \infty. \tag{6.11}$$

Proof Let

$$H(\lambda) := \int_0^\infty e^{-\lambda t}\, \mathbf{P}_0[\tau_0 > t]\, dt.$$

We may proceed as Fleischmann and Greven (1994), proof of Prop. 2.37 to obtain by a last exit decomposition

$$H(\lambda) = \frac{1-\lambda L_0(\lambda)}{\lambda(1+L_0(\lambda))}. \tag{6.12}$$

By Lemma 6.4 we get

$$H(\lambda) \sim \left(\frac{\gamma}{\alpha}\Gamma(1+\alpha)\right)^{-1} \lambda^{\alpha-1}, \qquad \lambda \to 0. \tag{6.13}$$

By the Tauberian theorem we get

$$\int_0^t \mathbf{P}_0[\tau_0 > s]\, ds \sim \left(\frac{\gamma}{\alpha}\Gamma(1+\alpha)\Gamma(2-\alpha)\right)^{-1} t^{1-\alpha}, \qquad t \to \infty. \tag{6.14}$$

Differentiating this formula w.r.t. t yields the claim. □

We will need the following comparison between delayed coalescing random walks $(\eta(t))$ and instantaneously coalescing random walks $(\widetilde{\eta}(t))$.

Lemma 6.6 *Uniformly in $\xi \in \Xi$ the following holds*

$$\mathbf{P}^{\{0,\xi\}}[\#\widetilde{\eta}(t) = \#\eta(t)] \overset{t\to\infty}{\longrightarrow} 1.$$

Proof Denote by $\varepsilon(t)$ any quantity vanishing uniformly in ξ as $t \to \infty$. Fix $0 < \beta < (\log Nc)^{-1}$. Then

$$\begin{aligned} \mathbf{P}^{\{0,\xi\}}[\#\widetilde{\eta}(t) < \#\eta(t)] &= \mathbf{P}^{\{0,\xi\}}[\#\widetilde{\eta}(t) = 1, \#\eta(t) = 2] \\ &\le \sum_{E\subset\Xi,\ \#E=2} \mathbf{P}^{\{0,\xi\}}[\widetilde{\eta}(t-\log t) = E, \#\widetilde{\eta}(t) = 1] \\ &\qquad + \mathbf{P}^{\{0,\xi\}}[\#\widetilde{\eta}(t-\log t) = 1, \#\eta(t) = 2]. \end{aligned} \tag{6.15}$$

By the recurrence of the random walk the last term is $\varepsilon(t)$. Thus by Corollary 6.3 the r.h.s. of (6.15) equals

$$\begin{aligned} &= \sum_{\substack{\xi_1,\xi_2\in\Xi \\ \|\xi_1-\xi_2\| \ge \beta\log t}} \mathbf{P}^{\{0,\xi\}}[\widetilde{\eta}(t-\log t) = \{\xi_1,\xi_2\},\ \#\widetilde{\eta}(t) = 1] + \varepsilon(t) \\ &= \sum_{\substack{\xi_1,\xi_2\in\Xi \\ \|\xi_1-\xi_2\| \ge \beta\log t}} \mathbf{P}^{\{0,\xi\}}[\widetilde{\eta}(t-\log t) = \{\xi_1,\xi_2\}] \cdot \mathbf{P}^{\{\xi_1,\xi_2\}}[\#\widetilde{\eta}(\log t) = 1] + \varepsilon(t) \\ &= \varepsilon(t). \end{aligned} \tag{6.16}$$

In the last two steps we have used the Markov property and the basic estimate on the hitting time of the origin. □

The case a critically recurrent

Consider now the case a critical. We adopt the notation from the preceding lemmas. Define a function $h :]0,1] \to [0,\infty[$ by

$$h(\lambda) = a_n \text{ for } \frac{1}{\lambda} = a_n N^n \tag{6.17}$$

(recall a_n from (2.6)) and linearly interpolated in the intervals $]a_n N^n, a_{n+1}N^{n+1}[$. Note that $h(\lambda)$ is slowly varying as $\lambda \to 0$.

Lemma 6.7

$$\mathbf{P}_0[\tau_0 > 1] \sim (h(1/t))^{-1}, \qquad t \to \infty. \tag{6.18}$$

Proof As was shown in the proof of Proposition 2.2 (i) and (ii) we have

$$L(\lambda/a_n N^n) \sim a_n, \qquad n \to \infty. \tag{6.19}$$

Thus

$$L(\lambda) \sim h(\lambda), \qquad \lambda \to 0. \tag{6.20}$$

We proceed as above to infer

$$H(\lambda) \sim \frac{1 - \lambda h(\lambda)}{\lambda(1 + h(\lambda))} \sim \frac{1}{\lambda h(\lambda)}, \qquad \lambda \to 0. \tag{6.21}$$

Apply the Tauberian theorem to complete the proof. □

Extended Duality Relation

For the proof of the proposition we will rely on a duality relation extending the basic duality. For $r, s \geq 0$ consider the following two particle system $\eta(r,s)$(introduced by Cox and Griffeath (1983)):

- Particle 1 stands still up to time $r - (r \wedge s)$ and then moves according to the random walk associated with the interaction kernel $a(\cdot,\cdot)$.
- Particle 2 stands still up to time $s - (r \wedge s)$ and then moves according to the random walk associated with the interaction kernel $a(\cdot,\cdot)$.
- After time $(r \wedge s)$ the particles are allowed to coalesce with delay whenever they occupy the same site.

In the same way we introduce $\widetilde{\eta}(r,s)$ where the coalescence is instantaneously.

By the duality lemma we infer immediately

$$\mathbf{E}^z[x_\zeta(r) x_\xi(s)] = \mathbf{E}^{\{\zeta,\xi\}}\left[z^{\eta(r,s)}\right], \quad z \in [0,1]^\Xi,\ \zeta,\xi \in \Xi. \tag{6.22}$$

By Lemma 6.6 this implies

$$\mathbf{E}^z[x_\zeta(r) x_\xi(s)] - \mathbf{E}^{\{\zeta,\xi\}}\left[z^{\widetilde{\eta}(r,s)}\right] \overset{|r-s|\to\infty}{\longrightarrow} 0. \tag{6.23}$$

We are now in the position to give the asymptotics of the covariance of $T_t(0)$ and $T_t(\xi)$.

Lemma 6.8 *Let $\theta \in [0,1]$ and $\mu \in \mathcal{M}_\theta$. Then the following holds*

$$\lim_{t\to\infty} t^{-2}\mathrm{Cov}^\mu[T_t(0), T_t(\xi)] = 4\theta(1-\theta)\cdot \lim_{t\to\infty} t^{-2}\int_0^t dv \int_0^{t-v} dw\, \mathbf{P}_0[\tau_0 > 2w]\int_v^{2v} ds\, p(s,0,\xi). \tag{6.24}$$

Proof

$$\begin{aligned}
t^{-2}\mathrm{Cov}^\mu[T_t(0), T_t(\xi)] &= t^{-2}\int_0^t ds \int_0^t dr\, \mathbf{E}^\mu[(x_0(r)-\theta)(x_\xi(s)-\theta)] \\
&= t^{-2}\int_0^t ds \int_0^t dr \int \mu(dz)\, \mathbf{E}^{\{0,\xi\}}\left[z^{\eta(r,s)}\right] - \theta^2 \\
&= t^{-2}\int_0^t ds \int_0^t dr\, 1\!\!1_{\{|r-s|>\log t\}} \int \mu(dz)\, \mathbf{E}^{\{0,\xi\}}\left[z^{\eta(r,s)}\right] - \theta^2 + \varepsilon(t) \\
&\overset{(6.23)}{=} t^{-2}\int_0^t ds \int_0^t dr\, 1\!\!1_{\{|r-s|>\log t\}} \int \mu(dz)\, \mathbf{E}^{\{0,\xi\}}\left[z^{\widetilde{\eta}(r,s)}\right] - \theta^2 + \varepsilon(t) \\
&= t^{-2}\int_0^t ds \int_0^t dr \int \mu(dz)\, \mathbf{E}^{\{0,\xi\}}\left[z^{\widetilde{\eta}(r,s)}\right] - \theta^2 + \varepsilon(t) \\
&= \theta(1-\theta)\cdot t^{-2}\int_0^t ds \int_0^t dr\, \mathbf{P}^{\{0,\xi\}}[\#\widetilde{\eta}(r,s) = 1] + \varepsilon(t).
\end{aligned}$$

In the last step we used the ergodic theorem. We may now proceed as in Cox and Griffeath (1983) (derivation of equation (2.2)) to infer the claim. □

We are now able the to give the proof of the proposition.

Proof of Proposition 6.1

Part (a). Assume that $a = a_c,\ c > 1$ is strongly recurrent. Substituting the results of Lemma 6.2 and Lemma 6.5 in (6.24) we obtain

$$\begin{aligned}
\lim_{t\to\infty} t^{-2}\mathrm{Var}^\mu[T_t(0)] &= 4\theta(1-\theta)\lim_{t\to\infty} t^{-2}\int_0^t dv \int_0^{t-v} dw\, \frac{(1-\alpha)2^{-\alpha}w^{-\alpha}}{\frac{\gamma}{\alpha}\Gamma(1+\alpha)\Gamma(2-\alpha)}\frac{a}{\alpha}(2^\alpha - 1)v^\alpha \\
&= 4\theta(1-\theta)\frac{1-2^{-\alpha}}{\Gamma(1+\alpha)\Gamma(2-\alpha)}\int_0^1 v^\alpha(1-v)^{1-\alpha}\, dv \\
&= 4\theta(1-\theta)\frac{1-2^{-\alpha}}{2},
\end{aligned}$$

where in the last step we have used Euler's identity for the β-integral

$$\int_0^1 v^\alpha(1-v)^{1-\alpha}dv = \frac{\Gamma(1+\alpha)\Gamma(2-\alpha)}{\Gamma(3)}.$$

Part (b). Assume now that a is critical. Apply Lemma 6.7 to (6.24) to obtain

$$\begin{aligned}
&\limsup_{t\to\infty} t^{-2}\mathrm{Cov}^\mu[T_t(0), T_t(\xi)] \qquad (6.25)\\
&= 4\theta(1-\theta)\cdot \limsup_{t\to\infty} t^{-1}h(1/t)^{-1}\int_0^t dv \int_v^{2v} ds\, p(s,0,\xi)
\end{aligned}$$

$$\begin{aligned} &\leq 2\theta(1-\theta)\limsup_{t\to\infty} t^{-1}h(1/t)^{-1}\int_0^t s\,p(s,0,0)\,ds \\ &= 0 \quad \text{by (2.12).} \end{aligned}$$

Indeed, let (c_n) be as in (2.12). Fix $\varepsilon \in]0,1[$ and take t large enough that $n = n(t)$ can be chosen such that

$$c_n N^n \leq \varepsilon t \leq t \leq a_n N^n. \tag{6.26}$$

Then by (2.12)

$$t^{-1}h(1/t)^{-1}\int_0^t s\,p(s,0,0)\,ds \leq \varepsilon + t^{-1}h(1/t)^{-1}a_n \int_{\varepsilon t}^t N^n p(s,0,0)\,ds \overset{t\to\infty}{\longrightarrow} \varepsilon. \tag{6.27}$$

□

Part III
The Branching Models

In this part we will investigate the long time behaviour of critical branching Brownian motion and (finite variance) super Brownian motion (the so-called Dawson-Watanabe process) on $\mathbb{R}^d$. These processes are known to be persistent if $d \geq 3$, that is there exist nontrivial equilibrium measures. If $d \leq 2$ they cluster, i.e. the resp. process converges to the 0 configuration while the surviving mass piles up in so-called clusters.

We study the spatial profile of the clusters in the "critical" dimension $d = 2$ via multiple space scale analysis. We will also investigate the long time behaviour of these models restricted to finite boxes in $d \geq 2$. On the way we develop coupling and comparison methods for spatial branching models.

1 Introduction

1.1 Background

For several interacting infinite particle and related models there is a dichotomy between stability (i.e. nontrivial equilibrium measures exist) and clustering depending on transience resp. recurrence of the interaction kernel. So many infinite particle systems with site space $\mathbb{Z}^d$ or $\mathbb{R}^d$ and finite variance interaction are stable if $d \geq 3$ and cluster if $d = 1, 2$. This is well known e.g. for the voter model, linearly interacting diffusions with compact state space, branching Brownian motion, Dawson-Watanabe process etc.

The dimension $d = 2$ is "critical" in the sense that the Green function of the interaction kernel grows only on logarithmic scale and is thus "nearly finite". In the critical dimension the phenomenon of "diffusive clustering" occurs. This means clusters grow at a randomly chosen algebraic scale of order t^α, $\alpha \in [0, 1/2]$. For many models the structure of the clusters in the critical dimension is known. The voter model in $\mathbb{Z}^2$ has been investigated by Cox and Griffeath (1986). "Critical dimension" linearly interacting diffusions with compact state space on the so-called hierarchical group have been studied by Fleischmann and Greven (1994), Dawson and Greven (1993a, 1993b), Dawson, Greven and Vaillancourt (1995) and in Part II of this work. The employed techniques to describe clusters cover scaling, re-normalisation and the so-called interaction chain.

Non-compact models such as super random walk on $\mathbb{Z}^d$ and linearly interacting Brownian motions labelled by $\mathbb{Z}^d$ have been treated by Winter (1995) and Kopietz (1995).

Clusters of branching Brownian motion have been been studied by Fleischman (1978) and Lee (1991). Lee has rather precise statements for the dimension dependent rate at which the height of clusters grows conditioned on (local) non-extinction (Thm. 2.4). Lee does however not treat the question of spatial extension and profile of the clusters. His results are obtained by studying sub- and super solutions of the partial differential equation determining the Laplace functional.

The main point of this part is to determine the spatial profile of the clusters of branching and super Brownian motion in dimension $d = 2$. Unlike Lee (1991) we will not condition on local non-extinction but follow a different route. The nivellation of the local extinction will be obtained by "blowing up" the initial configuration. The "blow-up" also enables us to to give a description of the finite system (considered next) in terms of the so-called *finite systems scheme* (introduced by Cox and Greven (1990)) that emphasises the similarities to other models.

In the theory of interacting particle systems a systematic treatment of the comparison of finite to infinite systems in high dimensions can be found in Cox and Greven (1990) and (1994b). The critical dimension voter model has been studied by Cox and Greven (1991). Comparison of finite to infinite systems of linearly interacting diffusions labelled by the hierarchical group in high and critical dimension can be found in Part II of this work. In this part we will also relate the behaviour of our branching processes to that of their finite versions, defined on d-dimensional tori, in both cases $d \geq 3$ and $d = 2$.

One aim of this part is to exhibit how the clustering phenomenon can be studied with *probabilistic tools*. Namely by techniques from the theory of infinite particle systems. These will be applied to both branching particle systems and super processes. In particular we rely on moment calculations and develop coupling and comparison techniques in Section 3. Thus our approach is completely different from Lee's (1991) and provides a more probabilistic understanding of these processes. Also our methods might be more easily adapted to related problems.

1.2 The Models

We only give a short heuristic description of the considered models. An extensive treatment can be found in Dawson (1977) and (1993) and in Fleischman (1978). Nevertheless we have to give the basic definitions for random measures first.

Basic Definitions for Random Measures

Let E be a locally compact polish space. By $\mathcal{B}(E)$ we denote the Borel σ-field on E. A measure μ on $\mathcal{B}(E)$ is called *locally finite* if $\mu(K) < \infty$ for all bounded sets $K \subset E$. The space

$$\mathcal{M}(E) = \{\text{locally finite measures on } E\} \tag{1.1}$$

is a polish space topologized by $\mu_n \to \mu$ iff $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for all $f : E \to \mathbb{R}$ continuous with compact support. The space of random measures $\mathcal{M}_1(\mathcal{M}(E))$ equipped with the weak topology (denoted by "$\Longrightarrow$") is also polish (see e.g. Kallenberg (1983)).

For a signed measure μ we denote by

$$\|\mu\| = \|\mu\|_{\mathrm{TV}} = \sup\{\mu(B) - \mu(E \setminus B) \,:\, B \in \mathcal{B}(E)\}$$

the total variation of μ. We denote by $\mathcal{M}_{\mathrm{f}}(E)$ the space of finite measures in $\mathcal{M}(E)$

$$\mathcal{M}_{\mathrm{f}}(E) = \{\mu \in \mathcal{M}(E) \,:\, \mu(E) < \infty\}. \tag{1.2}$$

The space of (non-negative) integer valued measure μ on $\mathcal{B}(E)$ will be denoted by

$$\mathcal{N}(E) = \{\mu \in \mathcal{M}(E) \,:\, \mu(A) \in \{0, 1, 2, \ldots, \infty\} \quad \forall\, A \in \mathcal{B}(E)\}. \tag{1.3}$$

The space of finite measures in $\mathcal{N}(E)$ is denoted by

$$\mathcal{N}_{\mathrm{f}}(E) = \{\mu \in \mathcal{N}(E) \,:\, \mu(E) < \infty\}. \tag{1.4}$$

Branching Brownian Motion

Let $(S_t)_{t\geq 0}$ be the semi group of a Feller process on E. We will consider a particle moving on E according to (S_t) having an exponential life time with mean $\frac{1}{c}$. At the time of death, with probability p_k there will be an offspring of k particles, all located at the parent's position. The probability function (p_k) is a basic data of the process. The offspring behaves as k independent copies of the one-particle system started at time zero. The process started in $x \in E$ will be denoted by $(\eta^x_t)_{t\geq 0}$. Its state space is $\mathcal{N}_{\mathrm{f}}(E)$.

For initial configuration $\eta_0 = \sum_{i=1}^{\infty} \delta_{x_i}$ (δ_x = Dirac-measure on x) in $\mathcal{N}(E)$ we define

$$\eta_t = \sum_{i=1}^{\infty} \eta^i_t \quad , \tag{1.5}$$

where $((\eta^i_t)_{t\geq 0}\,,\, i \in \mathbb{N})$ are independent copies of $(\eta^{x_i}_t)_{t\geq 0}$. In the case $p_0 = p_2 = \frac{1}{2}$ we will refer to (η_t) as the *critical binary branching process associated with* (S_t). One main object of consideration will be the critical binary branching Brownian motion on $\mathbb{R}^d$ (shorthand BBM($\mathbb{R}^d$)).

Dawson-Watanabe Process

Next we consider the short life time high density limit of binary branching processes. Let $x \in E$ and $N \in \mathbb{N}$. Let $(\eta_t^N)_{t\geq 0}$ be the branching process corresponding to $p_0 = p_2 = \frac{1}{2}$ with expected life time $\frac{1}{cN}$ and with initial state $\eta_0^N = N \cdot \delta_x$. It is well known that the *diffusion limit*

$$(\zeta_t^x)_{t\geq 0} = \mathrm{w} - \lim_{N\to\infty} \left(\frac{1}{N} \zeta_t^N \right)_{t\geq 0} \tag{1.6}$$

exists and is a Markov process with values in $\mathcal{M}_\mathrm{f}(E)$ (see Dawson (1993), Section 4.4ff).

The total population $\zeta_t^x(E)$ is known to be Feller's branching diffusion $(Z_{t/2})$. This is the diffusion on $[0, \infty[$ with generator

$$x \frac{\partial^2}{(\partial x)^2} \,. \tag{1.7}$$

Hence the finiteness of ζ_t^x is clear. Also $\mathbf{P}[\zeta_t^x(E) = 0] \stackrel{t\to\infty}{\longrightarrow} 1$ since (Z_t) is a martingale and 0 is an absorbing boundary point.

Let $((\zeta_t^x)_{t\geq 0}\,,\, x \in E)$ be independent copies of this process with $\zeta_0^x = \delta_x$. For $\mu \in \mathcal{M}(E)$ we can define $(\zeta_t)_{t\geq 0}$ with initial configuration $\zeta_0 = \mu$ by

$$\zeta_t = \int \zeta_t^x \mu(dx). \tag{1.8}$$

The process (ζ_t) has values in $\mathcal{M}(E)$ and will be called the *super process associated with* (S_t). Of particular interest will be super Brownian motion on $\mathbb{R}^d$ (short hand SBM($\mathbb{R}^d$)).

Another more analytic, though less intuitive, description is the following. For $\mu \in \mathcal{M}(E)$ and $\phi : E \to \mathbb{R}$ measurable and μ-integrable or non-negative we write

$$\langle \mu, \phi \rangle = \int \phi \, d\mu. \tag{1.9}$$

For $\phi : E \to [0, \infty[$ bounded and measurable with compact support and for $t \geq 0$ we define the (nonlinear) operator V_t by

$$V_t\phi = S_t\phi - \frac{1}{2} c \int_0^t S_{t-s}((V_s\phi)^2) ds. \tag{1.10}$$

Then (ζ_t) is defined by its *log-Laplace-semi group* (V_t), this is by the relation

$$\langle \zeta_0, V_t\phi \rangle = -\log \mathbf{E}[\exp(-\langle \zeta_t, \phi \rangle)]. \tag{1.11}$$

A path wise construction of (ζ_t) can be found in Le Gall (1991).

From the scaling properties of Brownian motion on $\mathbb{R}^d$ and Feller's diffusion (i.e. $\mathcal{L}^{\rho/\alpha}[\alpha Z_\beta] = \mathcal{L}^\rho[Z_{\alpha\beta}]$) it is clear that SBM($\mathbb{R}^d$) has the following *basic scaling property*: For $\alpha > 0$ and $\mu \in \mathcal{M}(\mathbb{R}^d)$ let $\mu'(\cdot) = \alpha\mu(\alpha^{-1/2} \cdot)$. Then

$$\mathcal{L}^{\mu'} \left[\alpha^{-1} \zeta_{\alpha t}(\alpha^{1/2} \cdot) \right] = \mathcal{L}^\mu[\zeta_t(\cdot)]. \tag{1.12}$$

In particular for $d = 2$ and $\mu = \lambda$ (Lebesgue measure on $\mathbb{R}^2$) this becomes

$$\mathcal{L}^{\lambda} \left[\alpha^{-1} \zeta_{\alpha t}(\alpha^{1/2} \cdot) \right] = \mathcal{L}^\lambda[\zeta_t(\cdot)]. \tag{1.13}$$

For simplicity we will hence forward only consider (the expected life time) $c^{-1} = 1$.

1.3 Basic Ergodic Theory

In the following we will state the results for BBM($\mathbb{R}^d$) and SBM($\mathbb{R}^d$) simultaneously. For convenience we will thus denote by $(\psi_t)_{t\geq 0}$ either BBM($\mathbb{R}^d$) or SBM($\mathbb{R}^d$). Also let for $\rho \geq 0$

$$M(\rho) = \begin{cases} \mathcal{H}(\rho) & \text{if } (\psi_t) \text{ is BBM}(\mathbb{R}^d) \\ \rho \cdot \lambda & \text{if } (\psi_t) \text{ is SBM}(\mathbb{R}^d) \end{cases}, \tag{1.14}$$

where λ is the (d-dimensional) Lebesgue measure and $\mathcal{H}(\rho)$ is a Poisson point process on $\mathbb{R}^d$ with intensity $\rho \cdot \lambda$.

It is well known (see Dawson (1977) and Fleischman (1978)) that if $d = 1$ or $d = 2$ then (ψ_t) clusters, i.e.

$$\mathcal{L}^{M(\rho)}[\psi_t] \overset{t\to\infty}{\Longrightarrow} \delta_{\mathbf{0}} \qquad \forall \rho \geq 0, \tag{1.15}$$

where $\delta_{\mathbf{0}}$ means the unit mass on $\mathbf{0} \in \mathcal{M}(\mathbb{R}^d)$.

For any $d \geq 3$ (ψ_t) is *persistent* (or stable). This means that there exists a family $(\nu_\rho\,,\, \rho \geq 0)\,,\ \nu_\rho \in \mathcal{M}_1(\mathcal{M}(\mathbb{R}^d))$ of invariant (under the dynamics) measures such that

$$\mathcal{L}^{M(\rho)}[\psi_t] \overset{t\to\infty}{\Longrightarrow} \nu_\rho. \tag{1.16}$$

Of course, the ν_ρ depend on whether (ψ_t) is BBM($\mathbb{R}^d$) or SBM($\mathbb{R}^d$). In the first case $m \in \mathcal{N}(E)$ $\nu_\rho(dm)$-a.s. The ν_ρ have the following properties. ν_ρ is translation invariant and ergodic with intensity ρ, this is

$$\int \langle m, \phi \rangle\, \nu_\rho(dm) = \rho \cdot \langle \lambda, \phi \rangle \tag{1.17}$$

for $\phi : \mathbb{R}^d \to [0, \infty[$ measurable. By the additivity property (1.5) the ν_ρ form a convolution semi group

$$\nu_{\rho+\sigma} = \nu_\rho * \nu_\sigma, \qquad \rho,\ \sigma \geq 0. \tag{1.18}$$

Hence any ν_ρ is infinitely divisible and thus allows a description via its canonical measure. For details and proofs see Gorostiza and Wakolbinger (1991) Thm. 2.2 for ψ_t BBM($\mathbb{R}^d$) and Dawson (1977) for SBM($\mathbb{R}^d$). For extension of the basic ergodic theory to more general branching mechanisms and motion semi group see Gorostiza, Roelly and Wakolbinger (1992). Extensions to initial configuration with infinite intensity or that are not translation invariant see Bramson, Cox and Greven (1993) and (1995) for the $d = 1, 2$ resp. $d \geq 3$ case for ψ_t BBM($\mathbb{R}^d$) and SBM($\mathbb{R}^d$).

2 Results

2.1 Cluster formation for $d = 2$

Since the branching mechanism has mean 1 the local extinction implies the existence of relatively small areas where more and more mass piles up. We call this phenomenon *clustering*. Our goal is to determine the spatial profile of the clusters. One a way to do so is to condition a test set B on being in a cluster. This precise statement for (ψ_t) BBM($\mathbb{R}^2$) is given by Fleischman (1978)

$$\frac{\log t}{8\pi} \mathbf{P}^{M(1)} \left[\psi_t(B) > \frac{\log t}{8\pi} |B| x \right] \overset{t\to\infty}{\longrightarrow} e^{-x} \qquad, x > 0, \tag{2.1}$$

where $B \in \mathcal{B}(\mathbb{R}^2)$. Roughly speaking, with probability $\frac{8\pi}{\log t}$ we see a cluster of "height" $\frac{\log t}{8\pi}$ - times an exponential mean 1 random variable. For BBM($\mathbb{R}^2$) Lee (1991) has a more precise statement (Thm 2.4) due to conditioning on $\eta_t(B) > 0$. Lee studies sub- and super-solutions of Kolmogorov's equation for the Laplace functional. His methods probably apply to SBM but it is still open whether the same is true for branching random walk on the lattice resp. linearly interacting Feller's diffusions (super random walk). This reflects the fact that difference equations are usually more difficult to treat than the related differential equation.

Our approach to describe the structure of clusters is based on two concepts.

(1) Blow-up

At time $t > 1$ we define

$$\widetilde{\psi}_t = \widetilde{\psi}_t^0 := \frac{8\pi}{\log t}\psi_t \tag{2.2}$$

with

$$\mathcal{L}[\psi_0] = \widetilde{M}(t) := M\left(\frac{\log t}{8\pi}\right). \tag{2.3}$$

This serves first to obtain a nontrivial limiting probability of local non extinction. Secondly the height of the clusters is scaled down to have a nontrivial limit.

(2) Spatial rescaling

For (ψ_t) BBM resp. SBM let $I = [0,1]$ resp. $I =]-\infty, 1]$. We fix $\alpha \in I$ and define $(\widetilde{\psi}_t^\alpha)$ by

$$\widetilde{\psi}_t^\alpha(B) := \mathcal{S}_{\alpha,t}\widetilde{\psi}_t(B) := t^{-\alpha}\widetilde{\psi}_t(t^{\alpha/2}B), \tag{2.4}$$

where $\mathcal{S}_{\alpha,t} : \mathcal{M}(\mathbb{R}^2) \to \mathcal{M}(\mathbb{R}^2),\ \mu(\cdot) \mapsto t^{-\alpha}\mu(t^{\alpha/2}\,\cdot)$. As above we let $\widetilde{\psi}_t = \widetilde{\psi}_t^0$. This is the right notion since clusters turn out to grow spatially as $t^{\alpha/2}$ for any $\alpha \in I$.

Remark: Note that by blowing up and rescaling we do not loose too much information on the family structure. This is because the blow-up is so smooth that by (2.1) in the limit $t \to \infty$ we get a Poisson mean 1 number of families in each bounded set $B \in \mathcal{B}(\mathbb{R}^2)$. On the other hand the spatial extension of a typical family is of order $t^{\alpha/2},\ \alpha < 1$ random. Hence the rescaling does not cause an overlap of the families. The blow-up also proves useful to give a description of the finite versions of our branching models that underlines the similarities to other models.

Now we are able to formulate the first theorem (recall that (Z_t) is Feller's branching diffusion defined in (1.7)).

Theorem 1 (Infinite System, $d = 2$)
Let (ψ_t) be either BBM($\mathbb{R}^2$) or SBM($\mathbb{R}^2$) and $I = [0,1]$ resp. $I =]-\infty, 1]$. Fix $\alpha \in I$. Then the following holds

$$\mathcal{L}^{\widetilde{M}(t)}[\widetilde{\psi}_t^\alpha] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}^1[Z_{1-\alpha}\cdot\lambda]. \tag{2.5}$$

Theorem 1 gives a first rough description of the profile of clusters. The averaging procedure induced by scaling however looses information about the spatial structure inside blocks of size $t^{\alpha/2}$.

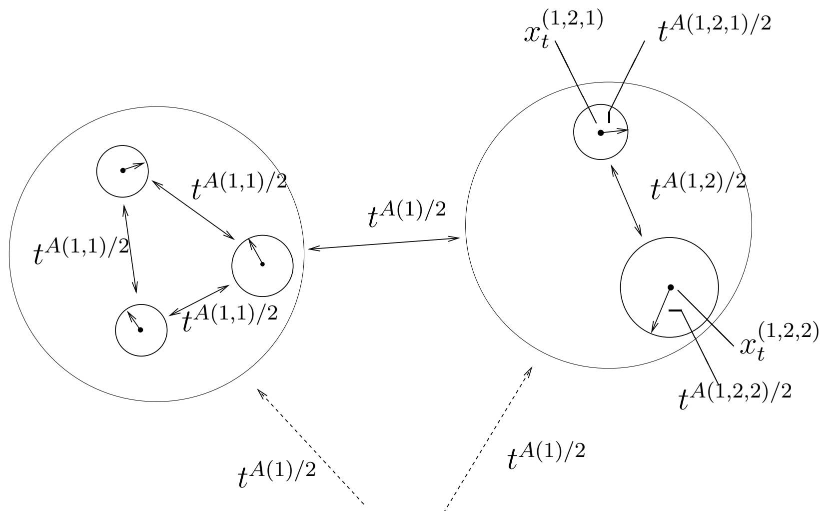

Figure 1. The points (dotted centers of the small circles) are grouped at distances growing at different scales $t^{A(\cdot)/2}$. The small circles represent the windows of observation which also grow at different scales.

The next aim is to give a more detailed description of the clusters via *multiple space scales.* That is, we want to look for different spatial scales on tuples of windows of observation (see Figure 1). To describe this properly on a formal level we introduce a rooted tree $\mathbb{T}$ (see Figure 2) and a space scale A associated with it.

Tree. We give the following representation of a (rooted) tree $\mathbb{T}$. Let $\mathbb{T}$ be a finite set of finite sequences with values in $\mathbb{N}$. The root will be denoted by $\emptyset \in \mathbb{T}$. Let $e, f \in \mathbb{T}$, $e = (e_1, \ldots, e_m)$, $f = (f_1, \ldots, f_n)$ (possibly $m = 0$ or $n = 0$) and $l = \max\{k : (e_1, \ldots, e_k) = (f_1, \ldots, f_k)\} \vee 0$. We then define the minimum $e \wedge f$ of e and f by

$$e \wedge f := \begin{cases} \emptyset & \text{if} \quad l = 0 \\ (e_1, \ldots, e_l) & \text{if} \quad l > 0. \end{cases}$$

We will assume that $(e_1, \ldots, e_k) \in \mathbb{T} \,\forall\, k \le m$ whenever $(e_1, \ldots, e_m) \in \mathbb{T}$. In particular this implies $e \wedge f \in \mathbb{T} \,\forall\, e, f \in \mathbb{T}$. $\mathbb{T}$ allows a partial ordering by $e \le f$ if and only if $e = e \wedge f$. The set of maximal elements in $\mathbb{T}$ will be denoted by $\mathbb{T}^M$. Note that we do not exclude the case in which $\mathbb{T}$ is partially ordered, i.e. $\#\mathbb{T}^M = 1$. In order to avoid redundancy we will assume that $(e_1, \ldots, e_{m-1}, g) \in \mathbb{T}$ for $g = 1, \ldots, e_m$ whenever $(e_1, \ldots, e_m) \in \mathbb{T}$.

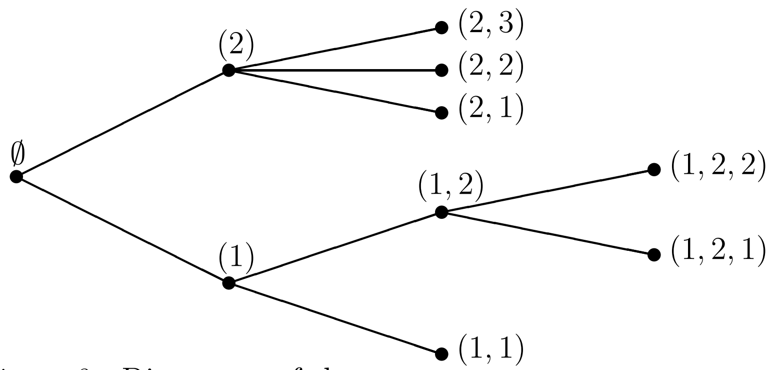

Figure 2. Diagramm of the tree
$\mathbb{T} = \{\emptyset, (1), (2), (1,1), (1,2), (1,2,1), (1,2,2), (2,1), (2,2), (2,3)\}$

Space scale. A pair $\mathbb{L} = (\mathbb{T}, A)$ consisting of a tree $\mathbb{T}$ and a strictly decreasing map

$$A : \mathbb{T} \to I$$

(recall that $I = [0,1]$ resp. $I =]-\infty, 1]$ in the case of BBM resp. SBM) will be called a *multiple space scale*. Given a multiple space scale $\mathbb{L} = (\mathbb{T}, A)$ we assume that $X = (x_t^e,\, e \in \mathbb{T},\, t \geq 0)$ is a family of points $x_t^e \in \mathbb{R}^2$ such that

$$\|x_t^e - x_t^f\| \approx t^{A(e \wedge f)/2}, \quad \text{as } t \to \infty.$$

By $a_t \approx b_t$ we mean $(\log a_t)/(\log b_t) \stackrel{t\to\infty}{\longrightarrow} 1$. We refer to X as to be $\mathbb{L}$-*spaced*. Our goal is to investigate the common distribution of (recall $\mathcal{S}_{\alpha,t}$ from (2.4))

$$(\mathcal{S}_{A(e),t}\mathcal{T}_{x_t^e}\widetilde{\psi}_t)_{e\in\mathbb{T}} \quad \text{as } t \to \infty,$$

where $\mathcal{T}_z : \mathcal{M}(\mathbb{R}^d) \to \mathcal{M}(\mathbb{R}^d)$ is the translation by z, $(\mathcal{T}_z\mu)(\cdot) = \mu(z + \cdot)$.

Feller tree. Let $(Z_t^e,\, e \in \mathbb{T})_{t\geq 0}$ be the following diffusion on $\mathbb{R}^{\mathbb{T}}$. Each $(Z_t^e)_{t\geq 0}$ is a Feller diffusion. For $e, f \in \mathbb{T}$ with $e \neq f$ we let $Z_t^e = Z_t^f$ for $t \in [0, 1 - A(e \wedge f)]$. For $t > 1 - A(e \wedge f)$ the evolutions of Z_t^e and Z_t^f shall be independent (see Figure 3).

Theorem 2 (Infinite System, Multiple Scale)
Let (ψ_t) be either BBM($\mathbb{R}^2$) or SBM($\mathbb{R}^2$). Then the following holds

(**a**)
$$\mathcal{L}^{\widetilde{M}(t)}\left[(\mathcal{S}_{A(e),t}\mathcal{T}_{x_t^e}\widetilde{\psi}_t)_{e\in\mathbb{T}}\right] \stackrel{t\to\infty}{\Longrightarrow} \mathcal{L}\left[\left(Z_{1-A(e)}^e \cdot \lambda\right)_{e\in\mathbb{T}}\right].$$

In particular for $\mathbb{T}$ linear

(**b**)
$$\mathcal{L}^{\widetilde{M}(t)}\left[(\widetilde{\psi}_t^\alpha(B))_{\alpha\in I}\right] \underset{\text{fdd}}{\stackrel{t\to\infty}{\Longrightarrow}} \mathcal{L}^1\left[|B| \cdot (Z_{1-\alpha})_{\alpha\in I}\right], \quad B \in \mathcal{B}(\mathbb{R}^d).$$

Remarks:

1. Since $Z_{1-A(e)}^e = Z_{1-A(e)}^f$ for $e \leq f$ it would suffice to define (Z_t^e) only for $e \in \mathbb{T}^M$.

2. In order to understand why Theorem 2 should be true we draw a time-space picture (see Figure 4). Consider a point $(x, t) \in \mathbb{R}^2 \times [0, \infty[$ in the "four space" (which is actually a "three space"). We want to investigate the events $C(x, t)$ that form the

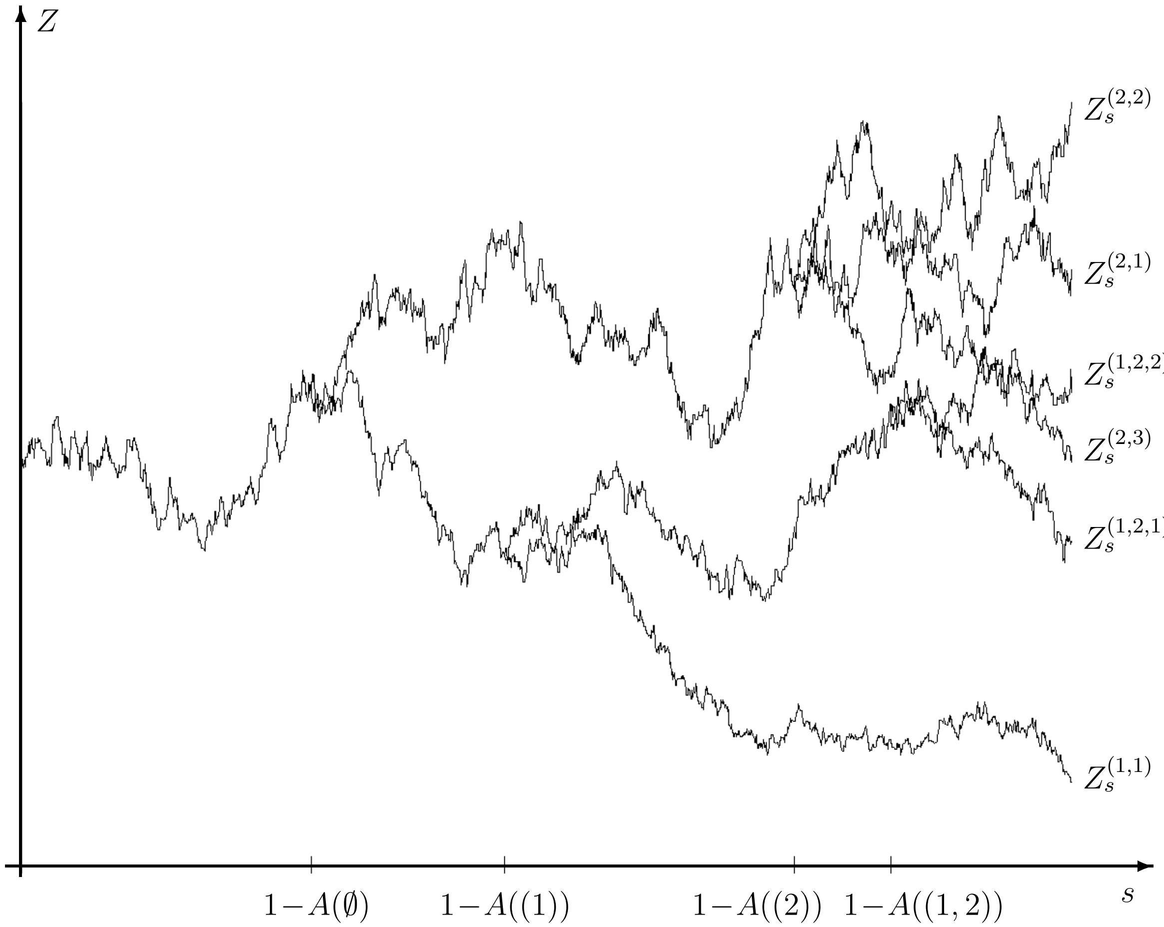

Figure 3. A sample of $(Z_s^e)_{s\geq 0}$, $e \in \mathbb{T}^M$ for $\mathbb{T} = \{\emptyset, (1), (2), (1,1), (1,2), (1,2,1), (1,2,2), (2,1), (2,2), (2,3)\}$

---

history of (x,t). Since Brownian motion at time s has range $\sim \sqrt{s}$ we may roughly set

$$C(x,t) = \{(u,s), \|u-x\| \leq (t-s)^{1/2},\ u \in \mathbb{R}^2,\ s \in [0,t]\}.$$

Now let for $\alpha \in [0,1]$

$$C_\alpha(x,t) = C(x,t) \cap (\mathbb{R}^2 \times \{t - t^\alpha\})$$

be the events at time $t-t^\alpha$ that may influence (x,t). Fix $\alpha \in [0,1]$ and let (x_t), $(y_t) \in \mathbb{R}^2$ be such that $\|x_t - y_t\| \sim t^{\alpha/2}$. Then for $\gamma < \alpha$ we have that $C_\gamma(x_t,t)$ and $C_\gamma(y_t,t)$ are (asymptotically) completely disjoint. For $\beta > \alpha$ we have that $C_\beta(x_t,t)$ and $C_\beta(y_t,t)$ (asymptotically) overlap completely. By the Markov property the the common history is contained in $C_\alpha(x_t,t) \approx C_\alpha(y_t,t)$. After time $t-t^\alpha$ the evolutions leading to (x_t,t) and (y_t,t) are independent.

We have to justify that the information contained in $C_\alpha(x_t,t) \approx C_\alpha(y_t,t)$ is sufficiently well described by the common value of $Z_{1-\alpha}$. Technically this done by showing that the distribution of mass is not "too inhomogeneous". A more philosophical point of view would be the following. At each scale of observation quasi-equilibria are exhibited that are determined by their density. Observation at different scales shows a certain self-similarity of those quasi-equilibria. This is reflected by the fact that the transition between scales is determined by a homogeneous Markov process.

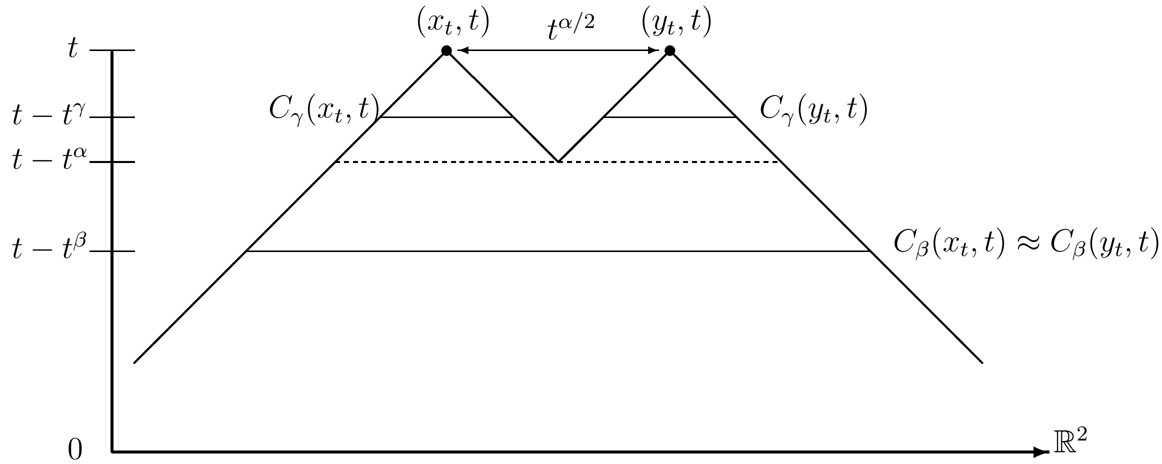

Figure 4. Historical cones for $\|x_t - y_t\| \sim t^{\alpha/2}$

2.2 Finite Systems, Stable Case

Computer simulations of particle systems evidently 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 idealisation 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?

We start with the definition of the finite versions of the d-dimensional BBM and SBM. Fix $d \in \mathbb{N}$ and let Λ_t^d for $t > 0$ the torus of size t, that is

$$\Lambda_t^d := \mathbb{R}^d/(t\mathbb{Z}^d). \tag{2.6}$$

We will regard Λ_t^d as the cube $[0, t[^d$ with periodic boundary conditions. Λ_t^d inherits the Brownian motion $(X_{t,s})_{s\geq 0}$ from $\mathbb{R}^d$. This is, $(X_{t,s})$ has transition densities

$$p_{t,s}(x, y) = \sum_{k\in\mathbb{Z}^d} p_s(x, y + tk), \tag{2.7}$$

where

$$p_s(x, y) = (2\pi s)^{-d/2} \exp\left(-\frac{\|x - y\|^2}{2s}\right) \tag{2.8}$$

is the transition density of d-dimensional Brownian motion. Finally denote by $M_t(\rho)$, $\mathcal{H}_t(\rho)$ etc. the restrictions of $M(\rho)$, $\mathcal{H}(\rho)$ etc. to Λ_t^d.

The objects of interest will be critical binary branching Brownian motion $(\eta_{t,s})_{s\geq 0}$ on Λ_t^d (shorthand BBM(Λ_t^d)) and super Brownian motion $(\zeta_{t,s})_{s\geq 0}$ on Λ_t^d (shorthand SBM(Λ_t^d)). Again let $(\psi_{t,s})_{s\geq 0}$ be either BBM(Λ_t^d) or SBM(Λ_t^d). The behaviour of the system is dictated by the *empirical population density* of the finite system

$$t^{-d}\psi_{t,s}(\Lambda_t^d).$$

Note that we obtain

$$\mathcal{L}^{M_t(\rho)}\left[t^{-d}\psi_{t,T(t)}(\Lambda_t^d)\right] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}^\rho[Z_{\sigma/2}], \tag{2.9}$$

if the observation time $T(t)$ satisfies

$$t^{-d}T(t) \stackrel{t\to\infty}{\longrightarrow} \sigma \qquad , \quad \sigma \in [0,\infty]. \tag{2.10}$$

The idea of how to describe stable, i.e. $d \geq 3$, finite systems is suggested by Cox and Greven (1990) and (1994b): The system is dominated by the macroscopic variable of the empirical population density. Roughly speaking it relaxes to the "equilibrium state" ν_θ with intensity θ, given that the empirical population density is θ. This relaxation takes place faster than the fluctuation of the empirical population density.

Thus by (2.9) t^d is the right time scale to look at the finite system. At this scale the empirical population density becomes random.

With these heuristics we are prepared for (recall ν_ρ from (1.16))

Theorem 3 (Finite System, Stable Case)
Let $d \geq 3$ and $(\psi_{t,s})_{s\geq 0}$ be either BBM(Λ_t^d) or SBM(Λ_t^d). Fix $\sigma \in [0,\infty]$ and $T(t)$ such that $t^{-d}T(t) \stackrel{t\to\infty}{\longrightarrow} \sigma$. Then the following holds

$$\mathcal{L}^{M_t(\rho)}\left[\psi_{t,T(t)}\right] \stackrel{t\to\infty}{\Longrightarrow} \int_0^1 \mathbf{P}^\rho[Z_{\sigma/2} \in d\theta]\nu_\theta. \tag{2.11}$$

2.3 Finite Systems, Critical Dimension

In dimension $d = 2$ we have to modify the ideas developed above in the fashion of rescaling presented in Subsection 2.1.

Fix $\alpha \in I$ and let for $s, t > 1$ (recall (2.4))

$$\widetilde{\psi}_{t,s}^\alpha(B) = \frac{8\pi}{\log s} s^{-\alpha}\psi_{t,s}\left((s^{\alpha/2}B) \cap \Lambda_t^2\right) \qquad , \quad B \in \mathcal{B}(\mathbb{R}^2). \tag{2.12}$$

Denote by $\widetilde{M_t}(s)$ the restriction of $\widetilde{M}(s)$ to Λ_t^2. Then

$$\mathcal{L}^{\widetilde{M}_t(T(t))}\left[\widetilde{\psi}_{t,T(t)}(\Lambda_t^2)\right] \stackrel{t\to\infty}{\Longrightarrow} \mathcal{L}^1[Z_{4\pi\sigma}], \tag{2.13}$$

if the observation time $T(t)$ satisfies

$$\frac{T(t)}{\beta(t)} \stackrel{t\to\infty}{\longrightarrow} \sigma \qquad , \quad \sigma \in [0,\infty]. \tag{2.14}$$

Here

$$\beta(t) = t^2 \log t. \tag{2.15}$$

It is due to the blow-up that $t^2 \log t$ is the right time scale to be used in the critical dimension. Many models in the critical dimension show a behaviour similar to (2.13). Namely linearly interacting diffusions with compact state space (Fisher-Wright, Fleming-Viot etc.), the voter model, etc. Interacting diffusions have been investigated in the "critical dimension" on the so-called hierarchical group by Fleischmann and Greven (1994), Dawson and Greven (1993a, 1993b), Dawson, Greven and Vaillancourt (1995) and in Part II of this work. Cox (1989) and Cox and Greven (1991) treat the voter model on $\mathbb{Z}^2$. The point seems to be that the Greens function of the interaction kernel is growing so slowly that taking the block averages is asymptotically the same as re-normalisation. Thus the role of the limiting diffusion (here the Feller diffusion in (2.13)) is played by the fixed point of the re-normalisation (see also Baillon et al. (1995)). The appropriate time scale in these models is the volume of the finite box times the recurrent potential kernel of the interaction kernel, maximised over the resp. finite box. For an extensive treatment of this latter point see Part II, Theorem 1.

Having in mind the proceeding of Subsection 2.1 the finite versions of Theorem 1 and 2 are easy to guess.

Theorem 4 (Finite System, $d = 2$)
Let $(\psi_{t,s})_{t,s}$ be either $BBM(\Lambda_t^2)$ or $SBM(\Lambda_t^2)$ and $I = [0,1]$ resp. $]-\infty, 1]$. Fix $\sigma \in [0,\infty]$ and $T(t)$ such that $T(t)/\beta(t) \overset{t\to\infty}{\longrightarrow} \sigma$. Then the following holds

$$\mathcal{L}^{\widetilde{M}_t(T(t))}\left(\widetilde{\psi}^{\alpha}_{t,T(t)}\right) \overset{t\to\infty}{\Longrightarrow} \int_0^\infty \mathbf{P}^1[Z_{2\pi\sigma} \in d\rho]\mathcal{L}^{\rho}[Z_{1-\alpha}] = \mathcal{L}^1[Z_{2\pi\sigma+1-\alpha}], \quad \alpha \in I. \tag{2.16}$$

Remark: Cox and Greven (1991) suggested to study the asymptotics of occupation times for the related model of branching random walk on $\mathbb{Z}^2$. Note that our result is more detailed than a description of the occupation time in that a time average is not made.

Let $\mathbb{L} = (\mathbb{T}, A)$ be a multiple space scale and let $X = (x_t^e, e \in \mathbb{T}, t \geq 0)$ be $\mathbb{L}$-scaled.

Theorem 5 (Finite System, Multiple Scale)
Under the conditions of Theorem 4 the following holds

$$\text{(a)} \quad \mathcal{L}^{\widetilde{M}(t)}\left[\left(\mathcal{S}_{A(e),T(t)}\mathcal{T}_{x_t^e}\widetilde{\psi}_{t,T(t)}\right)_{e\in\mathbb{T}}\right] \quad \overset{t\to\infty}{\Longrightarrow} \quad \int_0^\infty \mathbf{P}^1[Z_{2\pi\sigma} \in d\rho]\,\mathcal{L}^{\rho}\left[\left(Z^e_{1-A(e)}\cdot\lambda\right)_{e\in\mathbb{T}}\right].$$

In particular for $\mathbb{T}$ linear

$$\text{(b)} \qquad \mathcal{L}^{\widetilde{M}_t(T(t))}\left[(\widetilde{\psi}^{\alpha}_{t,T(t)}(B))_{\alpha\in I}\right] \quad \underset{\text{fdd}}{\overset{t\to\infty}{\Longrightarrow}} \quad \mathcal{L}^1\Big[|B|\cdot(Z_{2\pi\sigma+1-\alpha})_{\alpha\in I}\Big], \quad B \in \mathcal{B}(\mathbb{R}^d).$$

2.4 Outline

The rest of Part III is organised as follows. In Section 3 we will collect some tools needed later. This includes moment formulas, coupling techniques and comparison techniques. In Section 4 we prepare for the proof of Theorem 1 with an, admittedly, rather tedious moment calculation. Theorem 1 will be proved in Section 5. There we also apply the refined coupling methods in order to prove Theorem 2. In Section 6 the finite version theorems are proved with the comparison techniques from Section 3.

3 Basic Tools

In this section we develop the following tools for the investigation of the long time behaviour of our branching processes:

- We give a general basic *coupling* lemma and then give its applications to the special setting of an underlying Brownian motion. A further refinement will be obtained by the so-called local coupling (Lemma 3.5). This is the main result of this section. It serves to speed up the coupling. Hence it overcomes the difficulty that the subsequently given comparison technique works only for times $L(t)$ of order $L(t) \ll t^2$.

- A simple *comparison* technique
- $n-$th *moment* (recursion) *formulas*

For logical reasons we start with the presentation of the moment formulas.

3.1 Moment Formulas

Let E be either $\mathbb{R}^d$ or Λ_t^d. We will develop recursion formulas for the moments of BBM(E) and SBM(E).

We start with $(\eta_t)_{t\geq 0}$ BBM(E).

Lemma 3.1 (Moment Formula, BBM) *Let $(\eta_t)_{t\geq 0}$ be a BBM(E), where E is Λ_t^d resp. $\mathbb{R}^d$. Denote by $(S_t)_{t\geq 0}$ the semigroup of Brownian motion on E.*

(a) *For $n \in \mathbb{N}$, $x \in E$ and $\phi : E \to \mathbb{R}$ measurable and bounded or non-negative the n-th moment fulfils the following recursion formula*

$$\mathbf{E}^x[\langle \eta_t, \phi\rangle^n] = \langle \delta_x, S_t\phi\rangle + \frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\int_0^t S_{t-s}\left(\mathbf{E}^{\cdot}[\langle \eta_s, \phi\rangle^k]\mathbf{E}^{\cdot}[\langle \eta_s, \phi\rangle^{n-k}]\right)(x)\,ds. \quad (3.1)$$

In particular the first and second moments are

$$\mathbf{E}^x\left[\langle \eta_t, \phi\rangle\right] = \langle \delta_x, S_t\phi\rangle \quad (3.2)$$

$$\mathbf{E}^x\left[\langle \eta_t, \phi\rangle^2\right] = \langle \delta_x, S_t\phi\rangle + \left\langle \delta_x, \int_0^t S_{t-s}((S_s\phi)^2)\,ds\right\rangle. \quad (3.3)$$

(b) *For $\mu \in \mathcal{N}_{\mathrm{f}}(E)$, or $\mu \in \mathcal{N}(E)$ and ϕ with compact support, the first and second moments are*

$$\mathbf{E}^\mu\left[\langle \eta_t, \phi\rangle\right] = \langle \mu, S_t\phi\rangle \quad (3.4)$$

$$\mathbf{E}^\mu\left[\langle \eta_t, \phi\rangle^2\right] = \langle \mu, S_t\phi\rangle^2 + \left\langle \mu, \int_0^t S_{t-s}((S_s\phi)^2)\,ds\right\rangle + \left\langle \mu, S_t(\phi^2) - (S_t\phi)^2\right\rangle. \quad (3.5)$$

Proof For $f : \mathcal{N}_{\mathrm{f}} \to \mathbb{R}$ in the domain of the generator of BBM($\mathbb{R}^d$) $f(\eta_t)$ fulfils the following Kolmogorov backward equation

$$\frac{\partial}{\partial t}\mathbf{E}^{\delta_x}\left[f(\eta_t)\right] = \frac{1}{2}\Delta\mathbf{E}^{\delta_x}\left[f(\eta_t)\right] + \frac{1}{2}\mathbf{E}^{2\delta_x}\left[f(\eta_t)\right] + \frac{1}{2}\mathbf{E}^{\mathbf{0}}\left[f(\eta_t)\right] - \mathbf{E}^{\delta_x}\left[f(\eta_t)\right], \quad (3.6)$$

where Δ denotes the Laplace operator with respect to x and $\mathbf{0} \in \mathcal{N}_{\mathrm{f}}(E)$ means the zero measure. (Here and in the sequel we use $\mathbf{E}^x$ for $\mathbf{E}^{\delta_x}$ to avoid double subscripts.) In particular for $\phi : E \to [0, \infty[$ twice continuously differentiable, $n \in \mathbb{N}$ and $f(\mu) = \langle \mu, \phi\rangle^n$ equation (3.6) becomes (using the independence of the particles)

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\Delta\right)\mathbf{E}^x\left[\langle \eta_t, \phi\rangle^n\right] = \frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\mathbf{E}^x[\langle \eta_t, \phi\rangle^k]\mathbf{E}^x[\langle \eta_t, \phi\rangle^{n-k}]. \quad (3.7)$$

Integrating this yields (3.1). By an approximation argument (3.7) holds for $\phi : E \to \mathbb{R}$ measurable and bounded or non-negative.

For the part (b) note that by the independence of the particles

$$\mathbf{E}^\mu[\langle \eta_t, \phi\rangle^2] = \langle \mu, S_t\phi\rangle + \int \mu(dx)\mathrm{Var}^x[\langle \eta_t, \phi\rangle] \quad (3.8)$$

and use part (a). □

We continue with a moment recursion formula for SBM(E).

Lemma 3.2 (Moment formula, SBM)
Let $(\zeta_t)_{t\geq 0}$ be a SBM(E), where E is Λ_t^d resp. $\mathbb{R}^d$. Recall that $(S_t)_{t\geq 0}$ is the semigroup of Brownian motion on E. Let $\phi : E \to \mathbb{R}$ be bounded, measurable and with compact support and be $\mu \in \mathcal{M}(E)$. Then for $t \geq 0$ and $n \in \mathbb{N}$

$$\mathbf{E}^\mu[\langle \zeta_t, \phi\rangle^n] = \sum_{k=0}^{n-1} \binom{n-1}{k} \langle \mu, u^{(n-k)}(t)\rangle \, \mathbf{E}^\mu[\langle \zeta_t, \phi\rangle^k], \tag{3.9}$$

where $u^{(n)}(t) : \mathbb{R}^d \to \mathbb{R}$ is defined by

$$u^{(n)}(t) = \begin{cases} S_t\phi & , \quad n = 1 \\ \frac{1}{2}\sum_{k=1}^{n-1} \binom{n}{k} \int_0^t S_{t-s}\left(u^{(k)}(s)u^{(n-k)}(s)\right) ds & , \quad n \geq 2. \end{cases} \tag{3.10}$$

In particular the first and second moments are

$$\mathbf{E}^\mu[\langle \zeta_t, \phi\rangle] = \langle \mu, S_t\phi\rangle \tag{3.11}$$

$$\mathbf{E}^\mu[\langle \zeta_t, \phi\rangle^2] = \langle \mu, S_t\phi\rangle^2 + \left\langle \mu, \int_0^t S_{t-s}((S_s\phi)^2)\, ds \right\rangle. \tag{3.12}$$

Note that the first moment coincides with that of BBM while the second moment of BBM is greater than that of SBM. This reflects the fact that the "motion part" of SBM is deterministic while that of BBM is random.

The result and the idea of the proof can be found in Dawson (1993), Lemma 4.7.1. Unfortunately there are some misprints. So we give the proof in detail.

Proof Recall from (1.11) that (V_t) is the log-Laplace-semigroup of (ζ_t). Also recall that we assumed $c = 1$ in (1.10).

For $\theta \geq 0$ and $n \in \mathbb{N}$ let

$$u^{(n)}(t, \theta) = (-1)^{n-1} \frac{\partial^n}{(\partial\theta)^n} V_t(\theta\phi) \tag{3.13}$$

and

$$u^{(0)}(t, \theta) = -V_t(\theta\phi).$$

We can calculate $u^{(n)}(t, \theta)$ recursively with (1.10)

$$u^{(n)}(t, \theta) = \begin{cases} S_t\phi & , \quad n = 1 \\ \frac{1}{2}c \int_0^t S_{t-s}\left(\sum_{k=0}^{n} \binom{n}{k} u^{(k)}(s, \theta)u^{(n-k)}(s, \theta)\right) ds & , \quad n \geq 2. \end{cases} \tag{3.14}$$

Differentiating (1.11) w.r.t. θ yields

$$\langle \mu, u^{(1)}(t, \theta)\rangle \, \mathbf{E}^\mu \left[\langle \zeta_t, \phi\rangle \exp(-\theta\langle \zeta_t, \phi\rangle)\right] = \mathbf{E}^\mu \left[\exp(-\theta\langle \zeta_t, \phi\rangle)\right]. \tag{3.15}$$

Differentiate equation (3.15) $(n-1)$-times w.r.t. θ to obtain

$$\mathbf{E}^\mu \left[\langle \zeta_t, \phi\rangle^n \exp(-\theta\langle \zeta_t, \phi\rangle)\right] = \sum_{k=0}^{n-1} \binom{n-1}{k} \langle \mu, u^{(n-k)}(t, \theta)\rangle \, \mathbf{E}^\mu \left[\langle \zeta_t, \phi\rangle^k \exp(-\theta\langle \zeta_t, \phi\rangle)\right]. \tag{3.16}$$

Evaluating (3.16) at $\theta = 0$ yields the assertion. □

3.2 Coupling

In this subsection we shall construct two different couplings for our processes, the so-called basic coupling lemma (Lemma 3.3) and the local coupling (Lemma 3.5). On the way we recall in Lemma 3.4 the usual coupling for Brownian motions. We start explaining the notion of coupling in general.

Let $(S_t)_{t\geq 0}$ be the semigroup of a Feller process on the polish space E. By a *coupling* we mean a bivariate Feller process $(X_t, Y_t)_{t\geq 0}$ such that (X_t) and (Y_t) are each copies of a Feller process with semigroup (S_t). These copies need not be independent and in general will not. Note this definition is more general than the usual. In particular our coupling needs not be successful. In fact, we will use different notions of the "success" of a coupling.

Let

$$\tau = \inf\{t \geq 0 : X_t = Y_t\}. \tag{3.17}$$

We say that the coupling is *successful* for $(x, y) \in E \times E$ if

$$\mathbf{P}^{(x,y)}[\tau < \infty] = 1 \tag{3.18}$$

and

$$\mathbf{P}^{(x,y)}[\{X_t \neq Y_t\} \cap \{\tau < t\}] = 0 \qquad \forall\, t \geq 0. \tag{3.19}$$

We come to the first coupling (basic coupling). It deals with the coupling of two deterministic initial configurations μ^1 and μ^2.

Let $\mu^1, \mu^2 \in \mathcal{M}_\mathrm{f}(E)$ and let H be a non-negative random variable such that

$$\mathcal{L}^{(x,y)}[\tau] \leq \mathcal{L}[H] \text{ stochastically for } \mu^1 \otimes \mu^2\text{-almost all } (x, y) \tag{3.20}$$

and assume that (3.19) holds. For $A \in \mathcal{B}(E)$ let

$$C_t(A) = \sup\{(S_t \mathbb{1}_A)(x),\, x \in \mathrm{supp}(\mu^1 + \mu^2)\}.$$

Let $(\gamma_t^1)_{t\geq 0}$ and $(\gamma_t^2)_{t\geq 0}$ be binary branching processes resp. super processes associated with (S_t). In the former case we will also assume that $\mu^1, \mu^2 \in \mathcal{N}_\mathrm{f}(E)$.

Lemma 3.3 (Basic Coupling)
There exists a coupling $(\gamma_t^1, \gamma_t^2)_{t\geq 0}$ with $\gamma_0 = (\mu^1, \mu^2)$ that is successful in the sense that

$$\mathbf{E}\left[\left\|(\gamma_t^1 - \gamma_t^2)\Big|_A\right\|\right] \leq C_t(A) \cdot \Big|\|\mu^1\| - \|\mu^2\|\Big| + 2\min(\|\mu^1\|, \|\mu^2\|) \cdot \mathbf{P}[H > t]. \tag{3.21}$$

In particular for $\|\mu^1\| = \|\mu^2\|$

$$\mathbf{E}\left[\|(\gamma_t^1 - \gamma_t^2)\|\right] \leq 2\|\mu^1\| \cdot \mathbf{P}[H > t]. \tag{3.22}$$

Proof W.l.o.g. we may assume $\|\mu^1\| \leq \|\mu^2\|$. We make the decomposition

$$\mu^2 = \bar{\mu}^2 + \widetilde{\mu}^2$$

with $\|\bar{\mu}^2\| = \|\mu^1\|$. Then (3.20) holds with μ^2 replaced by either $\bar{\mu}^2$ or $\widetilde{\mu}^2$. It is clear (by the first moment formulas of the previous subsection) that (3.21) holds for any coupling

$\widetilde{\gamma}_t = (\widetilde{\gamma}_t^1, \widetilde{\gamma}_t^2)$ with $\widetilde{\gamma}_0 = (0, \widetilde{\mu}^2)$. Thus if we can show (3.22) for $(\bar{\gamma}_t)$ with $\bar{\gamma}_0 = (\mu^1, \bar{\mu}^2)$ we are done by setting $\gamma_t^i = \bar{\gamma}_t^i + \widetilde{\gamma}_t^i,\ i = 1, 2$.

Thus we will now assume $\|\mu^1\| = \|\mu^2\|$. Let $\mu \in \mathcal{M}_{\mathrm{f}}(E \times E)$ (resp. $\mu \in \mathcal{N}_{\mathrm{f}}(E \times E)$) with marginals $\mu^1(\cdot) = \mu(\cdot \times E)$ and $\mu^2(\cdot) = \mu(E \times \cdot)$. Let $(X_t, Y_t)_{t \geq 0}$ and τ be as above. Then we have by assumption

$$\mathbf{P}^{(x,y)}[X_t \neq Y_t] \leq \mathbf{P}[H > t] \qquad \text{for } \mu\text{-almost all } (x, y). \tag{3.23}$$

Define $(\gamma_t)_{t \geq 0}$ to be the critical branching (resp. super) process on $E \times E$ associated with the bivariate process $(X_t, Y_t)_{t \geq 0}$ on $E \times E$. For $t \geq 0$ we have that γ_t is in $\mathcal{M}_{\mathrm{f}}(E \times E)$ resp. $\mathcal{N}_{\mathrm{f}}(E \times E)$ almost surely. Let $\gamma_t^1(\cdot) = \gamma_t(\cdot \times E)$ and $\gamma_t^2(\cdot) = \gamma_t(E \times \cdot)$ be its marginals. Since the branching mechanism is spatially homogeneous it is clear that $(\gamma_t^1)_{t \geq 0}$ and $(\gamma_t^2)_{t \geq 0}$ are critical branching (resp. super) processes associated with (X_t) resp. (Y_t). Thus (γ_t^1) and (γ_t^2) are both associated with (S_t).

Denote by $D = \{(x, x) : x \in E\}$ the diagonal in $E \times E$. Then

$$\mathbf{E}^{\mu}\left[\left\|\gamma_t^1 - \gamma_t^2\right\|\right] \leq \mathbf{E}^{\mu}[\gamma_t((E \times E) \setminus D)] \leq \|\mu\| \cdot \mathbf{P}[H > t]. \tag{3.24}$$

□

We come back to the special situation $E = \mathbb{R}^d$ or $E = \Lambda_t^d$ and $(S_s)_{s \geq 0}$ the semigroup of Brownian motion on E. In this case there exists a successful coupling:

Lemma 3.4 *Let E be either Λ_t^d or $\mathbb{R}^d$ and let $R > 0$. For $x, y \in E$ with $\|x - y\| \leq R$ there exists a coupling $(W_s^1, W_s^2)_{s \geq 0}$ for the (standard) Brownian motion on E such that*

$$\mathbf{P}^{(x,y)}\left[W_s^1 \neq W_s^2\right] \leq \sqrt{\frac{1}{\pi}} R \cdot s^{-1/2}. \tag{3.25}$$

Proof We may assume $E = \mathbb{R}^d$ since on Λ_t^d the coupling works even better. By translation and orthogonal transformation we may also assume $x = 0$ and $y = (r, 0, \ldots, 0)$ with $r = \|x - y\| \leq R$.

If $d \geq 2$ we let

$$W_s^i = (Y_s^i, Z_s), \qquad i = 1, 2. \tag{3.26}$$

Here $(Z_s)_{s \geq 0}$ is a Brownian motion on $\mathbb{R}^{d-1}$ with $Z_0 = 0$. $(Y_s^1)_{s \geq 0}$ and $(Y_s^2)_{s \geq 0}$ are Brownian motions on $\mathbb{R}$ that move independently until they first meet and then move together. The initial points are $Y_0^1 = 0$ and $Y_0^2 = r$. In the case $d = 1$ we simply let $(W_s^i) = (Y_s^i),\ i = 1, 2$.

Let $H = \frac{1}{2} \inf\{s \geq 0 : Y_s^2 = 0\}$. Then (since $Y_s^2 - Y_s^1$ is a Brownian motion running at double speed)

$$\mathcal{L}[\inf\{s \geq 0 : W_s^1 = W_s^2\}] = \mathcal{L}^r[H]. \tag{3.27}$$

By the reflection principle

$$\mathbf{P}^r[H > s] = \sqrt{\frac{2}{\pi}} \int_0^{r/\sqrt{2s}} e^{-u^2/2}\, du \leq \sqrt{\frac{1}{\pi}} R s^{-1/2}. \tag{3.28}$$

□

The aim is now to couple the evolution of $(\psi_s)_{s\geq 0}$ started from two different (random) configurations. In the context of our problem one of those laws is only vaguely known since it will be the result of long-time evolution of a (ψ_s)-type process. The other will be better known and will typically be $M(\rho')$. Here the (random) value ρ' is obtained by some averaging over the first configuration. The details follow in the subsequent sections.

Since $\operatorname{supp}(\gamma^1+\gamma^2)$ will typically be too large to apply Lemma 3.4 directly we have to construct a local coupling. The idea is the following.

We start with a translation invariant initial configuration. Thus the support *is* large. In order to apply Lemma 3.4 successfully we divide E into boxes of length $R>0$. We do the coupling independently in each box according to Lemma 3.4. Finally we have to shift the pattern of boxes by a random offset $z\in[0,R[^d$ in order to obtain a translation invariant coupling.

Let $Q=Q(d\gamma^1,d\gamma^2)\in\mathcal{M}_1(\mathcal{M}(E)\times\mathcal{M}(E))$ be translation invariant. Fix $R>0$. In the case $E=\Lambda_t^d$ we will assume that $t/R=:N\in\mathbb{N}$.

Lemma 3.5 (Local Coupling)
There exists a (translation invariant) coupling $(\psi_s^1,\psi_s^2)_{s\geq 0}$ of BBM(E) resp. SBM(E) with

$$\mathcal{L}[(\psi_0^1,\psi_0^2)]=Q \tag{3.29}$$

and such that

$$\begin{aligned}&\mathbf{E}\left[\left\|(\psi_s^1-\psi_s^2)\Big|_A\right\|\right] \qquad (3.30)\\ &\quad\leq |A|\cdot\frac{1}{R^d}\left[\mathbf{E}\left[\left|(\psi_0^1-\psi_0^2)([0,R[^d)\right|\right]+\mathbf{E}\left[(\psi_0^1+\psi_0^2)([0,R[^d)\right]\cdot\sqrt{\frac{d}{\pi}}R\cdot s^{-1/2}\right].\end{aligned}$$

Proof Fix an initial configuration $(\mu^1,\mu^2)\in\mathcal{M}(E)\times\mathcal{M}(E)$. Let

$$C_k=kR+[0,R[^d \tag{3.31}$$

for $k\in\mathbb{Z}^d$ resp. $k\in\{0,\dots,N-1\}^d$. Let

$$\mu_k^i=\mu^i\mathbb{1}_{C_k},\qquad i=1,2 \text{ for each } k. \tag{3.32}$$

We want to use the independence in the branching systems to obtain a coupling $(\gamma_{k,s}^1,\gamma_{k,s}^2)_{s\geq 0}$ for μ_k^1 and μ_k^2 for each k separately. Fix k. We apply Lemma 3.3 and Lemma 3.4 with $A=E$ (note that two points in C_k have distance at most $R\sqrt{d}$) to get

$$\mathbf{E}^{(\mu_k^1,\mu_k^2)}\left[\left\|\gamma_{k,s}^1-\gamma_{k,s}^2\right\|\right]\leq\left|\|\mu_k^1\|-\|\mu_k^2\|\right|+2\min(\|\mu_k^1\|,\|\mu_k^2\|)\cdot\sqrt{\frac{d}{\pi}}R\cdot s^{-1/2}. \tag{3.33}$$

Integrating (3.33) with respect to $Q(d\mu^1,d\mu^2)$ and using translation invariance we get

$$\mathbf{E}\left[\left\|\gamma_{k,s}^1-\gamma_{k,s}^2\right\|\right]\leq\mathbf{E}\left[\left|(\psi_0^1-\psi_0^2)(C_0)\right|\right]+\mathbf{E}[(\psi_0^1+\psi_0^2)(C_0)]\cdot\sqrt{\frac{d}{\pi}}R\cdot s^{-1/2}=:\varepsilon. \tag{3.34}$$

If we let $\gamma_s^i=\sum_k\gamma_{k,s}^i$, $i=1,2$ then $\mathcal{L}[(\gamma_0^1,\gamma_0^2)]=Q$ and (by translation invariance)

$$\mathbf{E}\left[\left\|(\gamma_s^1-\gamma_s^2)\Big|_{C_k}\right\|\right]\leq\varepsilon\qquad\forall\,k. \tag{3.35}$$

In order to get a translation invariant coupling we pick $z \in C_0$ at random and shift the "grid" $R\mathbb{Z}^d$ by z: For $z \in C_0$ define $(\gamma^i_s(z))_{t\geq 0}$, $i = 1, 2$ as above with C_k replaced by $C_k(z) = z + C_k$. Let

$$\mathcal{L}[\psi^i_s] = \frac{1}{R^d} \int\limits_{C_0} \mathcal{L}[\gamma^i_s(z)]\, dz, \qquad i = 1, 2. \tag{3.36}$$

Then (ψ^1_s, ψ^2_s) is a coupling with the asserted properties: (3.29) holds because it holds for each $(\psi^1_0(z), \psi^2_0(z))$, $z \in C_0$. By construction $\mathbf{E}\left[\left\|(\psi^1_s - \psi^2_s)\Big|_B\right\|\right]$ is translation invariant on E as measure in B. Hence it is a multiple of the Lebesgue measure on E. By (3.35) its density is $\leq \varepsilon$.

□

Corollary 3.6 *Let $Q \in \mathcal{M}_1(\mathcal{M}(\Lambda^d_t) \times \mathcal{M}(\Lambda^d_t))$ resp. $\mathcal{M}_1(\mathcal{N}(\Lambda^d_t) \times \mathcal{N}(\Lambda^d_t))$ be translation invariant with*

$$\rho := t^{-d} \int \gamma^1(\Lambda^d_t)\, Q(d\gamma^1, d\gamma^2) < \infty. \tag{3.37}$$

Given γ^1 under $Q(d\gamma^1, d\gamma^2)$ the distribution of γ^2 shall be $M_t(\rho')$ with $\rho' := t^{-d}\gamma^1(\Lambda^d_t)$.

Let further $N \in \mathbb{N}$, $R = t/N$ and $\varepsilon > 0$ such that

$$\mathbf{E}[|\gamma^1(\Lambda^d_t) - N^d\gamma^1([0, R[^d)|] < \varepsilon t^d. \tag{3.38}$$

Then there exists a coupling $(\psi^1_{t,s}, \psi^2_{t,s})_{s\geq 0}$ of $BBM(\Lambda^d_t)$ resp. $SBM(\Lambda^d_t)$ with $\mathcal{L}[(\psi^1_{t,0}, \psi^2_{t,0})] = Q$ and such that for $B \in \mathcal{B}(\Lambda^d_t)$ and $s \geq 0$

$$\mathbf{E}\left[\left\|(\psi^1_{t,s} - \psi^2_{t,s})\Big|_B\right\|\right] \leq |B| \cdot \left[\varepsilon + 2\sqrt{\rho R^{-d}} + 2\sqrt{\frac{d}{\pi}}\,\rho R \cdot s^{-1/2}\right]. \tag{3.39}$$

Proof In the case of SBM clearly $\mathbf{E}[|(\psi^1_{t,0} - \psi^2_{t,0})([0, R[^d)|] \leq \varepsilon R^d$. Consider now the case of BBM. Note that for a Poisson random variable X with mean $\theta > 0$

$$\mathbf{E}[|X - \theta|] \leq \sqrt{\theta} + \frac{1}{\sqrt{\theta}}\mathrm{Var}[X] = 2\sqrt{\theta}. \tag{3.40}$$

By this and Jensen's inequality we obtain

$$\begin{aligned} \mathbf{E}\left[\left|(\psi^1_{t,0} - \psi^2_{t,0})([0, R[^d)\right|\right] &\leq \varepsilon R^d + \mathbf{E}\left[\left|\gamma^2([0, R[^d) - N^{-d}\gamma^1(\Lambda^d_t)\right|\right] \\ &\leq \varepsilon R^d + 2\mathbf{E}\left[\sqrt{N^{-d}\gamma^1(\Lambda^d_t)}\right] \\ &\leq \varepsilon R^d + 2\sqrt{\rho R^d}. \end{aligned} \tag{3.41}$$

Now apply Lemma 3.5. □

Corollary 3.7 *Let $S > R > 0$ and $E = \mathbb{R}^d$. Consider $(\psi^1_s)_{s\geq 0}$ $BBM(\mathbb{R}^d)$ resp. $SBM(\mathbb{R}^d)$. Assume that $\mathcal{L}[\psi^1_0]$ is translation invariant and that $\varepsilon, \delta > 0$ and $0 < \rho < \infty$ are chosen such that*

$$\begin{aligned} \mathbf{E}[\psi^1_0([0, 1[^d)] &= \rho \\ \mathbf{E}[|R^{-d}\psi^1_0([0, R[^d) - S^{-d}\psi^1_0([0, S[^d)|] &< \varepsilon \\ \mathbf{E}[|\psi^1_0([0, S[^d) - \psi^1_0(S(z + [0, 1[^d))|] &< \delta S^d \qquad \forall\, z \in [-1, 1]^d. \end{aligned}$$

Then there exists a coupling $(\psi_s^1, \psi_s^2)_{s\geq 0}$ *such that*

$$\mathcal{L}[\psi_0^2 \Big| \psi_0^1] = M(S^{-d}\psi_0^1([0,S[^d)) \tag{3.42}$$

and

$$\mathbf{E}[\|(\psi_s^1 - \psi_s^2)\Big|_B\|] \leq |B| \cdot \left[\varepsilon + 3\delta + d\, e^{-D^2/2s} + 2\sqrt{\rho R^{-d}} + 2\sqrt{\frac{d}{\pi}}\,\rho R s^{-1/2}\right], \tag{3.43}$$

where $B \in \mathcal{B}(\mathbb{R}^d)$, $B \subset [0,S[^d$ *and* $D = \operatorname{dist}(B, \mathbb{R}^d \setminus [0,S[^d)$.

Proof If the common distribution of ψ_0^1 and ψ_0^2 was translation invariant we could argue as in Corollary 3.6. However, in general it is not. So we have to work a little more. The aim is to construct a third process $(\psi_s^3)_{s\geq 0}$ such that $\mathcal{L}[\psi_0^1, \psi_0^3]$ is translation invariant while ψ_s^2 and ψ_s^3 are close. Here are the details.

Let for $\gamma \in \mathcal{M}(\mathbb{R}^d)$ and $z \in \mathbb{R}^d$

$$\Gamma(z,\gamma) = \sum_{k\in\mathbb{Z}^d} M(S^{-d}\gamma(S(z+k+[0,1[^d))) \cdot \mathbb{1}_{S(z+k+[0,1[^d)} \tag{3.44}$$

and

$$\mathcal{L}[\psi_0^3|\psi_0^1] = \int_{[0,1[^d} \Gamma(z,\psi_0^1)\, dz.$$

Then clearly (by a suitable coupling of the Poisson processes in (3.44) and (3.42) in the case of BBM) we can assume

$$\mathbf{E}[\|(\psi_0^3 - \psi_0^2)\Big|_A\|] \leq \delta|A|, \quad A \subset [0,S[^d, \tag{3.45}$$

which implies that we can couple (ψ_s^2) and $\psi_s^3)$ such that

$$\begin{aligned} \mathbf{E}[|(\psi_s^3 - \psi_s^2)(B)|] &\leq \delta|B| + 2\rho \int_{\mathbb{R}^d\setminus[0,S[^d} dx \int_B dy\, p_s(x,y) \\ &\leq |B|(\delta + 2\rho d\, e^{-D^2/2s}). \end{aligned} \tag{3.46}$$

(This coupling is done by defining three independent processes with initial configurations $\psi_0^2 \wedge \psi_0^3$, $(\psi_0^2 - \psi_0^3)^+$, $(\psi_0^2 - \psi_0^3)^-$.) As in (3.41) we get

$$\begin{aligned} \mathbf{E}[|(\psi_0^3 - \psi_0^1)([0,R[^d)|] &\leq \mathbf{E}\left[\left|\psi_0^3([0,R[^d) - \mathbf{E}[\psi_0^3([0,R[^d)]\right|\right] \\ &\quad + \mathbf{E}\left[|\psi_0^1([0,R[^d) - \mathbf{E}[\psi_0^3([0,R[^d)\Big|\psi_0^1]|\right] \\ &\leq 2\sqrt{\rho R^d} + (\varepsilon+\delta)R^d. \end{aligned} \tag{3.47}$$

Now apply Lemma 3.5 to (ψ_0^1, ψ_0^3). □

3.3 Comparison

In this subsection we compare the finite versions of our branching processes to their infinite versions. We show that the finite system is not "too far off" from its infinite counterpart if the time $L(t)$ of observation is not too large. Unfortunately "not too large" here means $L(t) \ll t^2$. Hence the obtained comparison result is not at all surprising. However, with the strong tool of local coupling this will be sufficient for our purposes.

Lemma 3.8 (Comparison) *Let $t \geq 0$ and $A \in \mathcal{B}(\Lambda_t^d)$ such that $D = \frac{1}{2}(t - \mathrm{diam}(A)) > 0$. There exist two BBM resp. SBM $(\psi_s^1)_{s\geq 0}$ on $\mathbb{R}^d$ and $(\psi_{t,s}^2)_{s\geq 0}$ on Λ_t^d on one probability space such that for $s > 0$*

$$\psi_0^1 = M(\rho) \qquad \textit{and} \qquad \psi_{t,0}^2 = M_t(\rho) \tag{3.48}$$

and

$$\mathbf{E}\left[\left|\psi_s^1(A) - \psi_{t,s}^2(A)\right|\right] \leq 2d \exp\left(-\frac{D^2}{2s}\right) \cdot \rho|A|\frac{\sqrt{s}}{D}. \tag{3.49}$$

In particular for a sequence $L(t) \ll t^2$ and $A_t = t^{\alpha/2}A$, $\alpha \in [0, 2[$ we get uniformly in $\rho > 0$

$$\frac{t^{-d\alpha/2}}{\rho|A|}\mathbf{E}\left[\left|\psi_{L(t)}^1(t^{\alpha/2}A) - \psi_{t,L(t)}^2(t^{\alpha/2}A)\right|\right] \overset{t\to\infty}{\longrightarrow} 0. \tag{3.50}$$

Proof W.l.o.g. we may assume that A is centered in Λ_t^d such that

$$\sup\{\|x-y\|,\, x \in A, y \in \mathbb{R}^d \setminus \Lambda_t^d\} \leq \frac{1}{2}(t - \mathrm{diam}(A)).$$

For $m \in \mathbb{Z}^d$ let $(\gamma_s^m)_{s\geq 0}$ be independent BBM($\mathbb{R}^d$) resp. SBM($\mathbb{R}^d$) with (independent) initial configurations

$$\mathcal{L}[\gamma_0^m] = M(\rho) \cdot \mathbb{I}_{t(m+[0,1[^d)}. \tag{3.51}$$

Let

$$\psi_s^1(\cdot) = \sum_{m\in\mathbb{Z}^d} \gamma_s^m(\cdot) \quad \text{and} \quad \psi_{t,s}^2(\cdot) = \sum_{m\in\mathbb{Z}^d} \gamma_s^0(mt + \cdot). \tag{3.52}$$

Then (ψ_s^1) and $(\psi_{t,s}^2)$ are as asserted and we have to show (3.49). By construction

$$\begin{aligned}
\mathbf{E}\left[\left|\psi_s^1(A) - \psi_{t,s}^2(A)\right|\right] &\leq \sum_{m\in\mathbb{Z}^d\setminus\{0\}} \mathbf{E}[\gamma_s^m(A)] + \mathbf{E}[\gamma_s^0(mt + A)] \\
&= 2 \sum_{m\in\mathbb{Z}^d\setminus\{0\}} \mathbf{E}\left[\gamma_s^0(mt + A)\right] \\
&= 2\rho \int_{\mathbb{R}^d\setminus\Lambda_t^d} dx \int_A dy\, p_s(x,y) \\
&\leq 2\rho|A|\mathbf{P}^0\Big[\|W_s\| \geq D\Big],
\end{aligned} \tag{3.53}$$

where $(W_s)_{s\geq 0}$ is a standard Brwonian motion on $\mathbb{R}^d$. The proof of (3.49) is now a standard estimate while (3.50) is an immediate consequence of (3.49). □

4 Moment Calculations in the Critical Dimension

In this section we give the asymptotics of the moments of BBM($\mathbb{R}^2$) and SBM($\mathbb{R}^2$). We will obtain bounds for the moments as well. These allow us to express the Laplace transform in terms of the moments in the next section.

Fix $B \in \mathcal{B}(\mathbb{R}^2)$ and let for $t \geq 0$

$$B_t = B_{\alpha,t} = t^{\alpha/2} B. \tag{4.1}$$

Now for $n \in \mathbb{N},\ x \in \mathbb{R}^2, s \geq 0$ and $t > 1$ let

$$m_n(x,s,t) \;=\; m_n(x,s,t,\alpha) = \mathbf{E}^x\left[(\psi_s(B_{\alpha,t}))^n\right] \tag{4.2}$$

$$\widetilde{m}_n(x,s,t) \;=\; \widetilde{m}_n(x,s,t,\alpha) = \frac{s}{(\log s)^{n-1}} t^{-n\alpha} \mathbf{E}^x\left[(\psi_s(B_{\alpha,t}))^n\right]. \tag{4.3}$$

By $\mathbf{E}^x$ we mean of course the expectation when the initial configuration is $\delta_x \in \mathcal{N}_{\mathrm{f}}(\mathbb{R}^2)$. Let (recall p_t from 2.8)

$$\varphi(x) = p_1(0,x) = \frac{1}{2\pi}\exp\left\{-\|x\|^2/2\right\}, \quad x \in \mathbb{R}^2. \tag{4.4}$$

Although an abuse of notation no problems will arise by suppressing the dimension in the notation. Fix $x \in \mathbb{R}^2$ and three non-negative sequences $(a_t) \downarrow 0, (b_t) \downarrow 0$ and $(c_t) \uparrow \infty$.

Lemma 4.1 *Let $B \in \mathcal{B}(\mathbb{R}^2)$ be bounded.*

(a) Uniformly in β such that $1 \geq \beta \geq \alpha$ and uniformly in the sequences $(x_t)_{t\geq 0}$ and $(s_t)_{t \geq 0}$ such that

$$\left|\frac{x_t}{\sqrt{s_t}} - x\right| < a_t \tag{4.5}$$

and

$$\left|\frac{\log s_t}{\log t} - \beta\right| < b_t \tag{4.6}$$

and such that $s_t > t^\alpha c_t$ the following holds

$$\lim_{t\to\infty} \widetilde{m}_n(x_t, s_t, t, \alpha) = \varphi(x)\left(1 - \frac{\alpha}{\beta}\right)^{n-1} \frac{|B|^n n!}{(8\pi)^{n-1}} \tag{4.7}$$

and

$$\lim_{t\to\infty} \frac{1}{s_t}\int_{\mathbb{R}^2} \widetilde{m}_n(y, s_t, t, \alpha)\, dy = \left(1 - \frac{\alpha}{\beta}\right)^{n-1} \frac{|B|^n n!}{(8\pi)^{n-1}}. \tag{4.8}$$

(b) There exists $\Gamma < \infty$ such that

$$\sup_{t: t \geq s_t \geq 3} \sup_{n\in\mathbb{N}} \frac{1}{n!\Gamma^n} \widetilde{m}_n(x_t, s_t, t, \alpha) < \infty \tag{4.9}$$

and

$$\sup_{t: t \geq s_t \geq 3} \sup_{n\in\mathbb{N}} \frac{1}{n!\Gamma^n} \frac{1}{s_t}\int_{\mathbb{R}^2} \widetilde{m}_n(y, s_t, t, \alpha)\, dy < \infty. \tag{4.10}$$

Proof Throughout this proof we will suppress the α where no ambiguities may occur.

Our main tool is the moment recursion formula for BBM($\mathbb{R}^d$)

$$\mathbf{E}^x[(\eta_s(A))^n] = \mathbf{E}^x[\eta_s(A)] \\ +\frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\int_0^s du \int_{\mathbb{R}^2} dy\, p_{s-u}(x,y)\mathbf{E}^y[(\eta_u(A))^k]\mathbf{E}^y[(\eta_u(A))^{n-k}] \quad \forall A \in \mathcal{B}(\mathbb{R}^2), \tag{4.11}$$

(this is (3.1) with $\phi = 1\!\!1_A$). In particular, for $A = B_{\alpha,t}$ equation (4.11) becomes

$$m_n(x,s,t) = m_1(x,s,t) + \frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\int_0^s du \int_{\mathbb{R}^2} dy\, p_{s-u}(x,y) m_k(y,u,t) m_{n-k}(y,u,t). \tag{4.12}$$

Compare this with the moment formula for SBM($\mathbb{R}^d$) given in Lemma 3.2. Since the leading terms coincide it suffices to prove the assertion for the case $(\psi_t) = (\eta_t)$ is BBM($\mathbb{R}^2$). Note that for the case (ψ_t) SBM also the existence of Γ with the asserted properties also follows easily from the existence in the case considered here.

We start with the proof of part (a). The proof follows an idea of Durrett (1979) (proof of Thm. 8.1). We proceed by induction over n using (4.12). To do so we cut the left and right side of the domain $[0, s_t]$ of integration. In the remaining term we may use the asymptotics (4.7) and (4.8). On the other hand the error terms from the truncation of the domain of integration will be estimated by the following bounds. These will be proved successively in the course of the induction.

We show the existence of constants C_n, D_n and E_n (depending on B) with

$$\sup_{\substack{t \geq s \geq u \geq 3 \\ y \in \mathbb{R}^2}} \frac{1}{u} \int_{\mathbb{R}^2} (s-u) p_{s-u}(y,z) \widetilde{m}_n(z,u,t) dz \leq C_n, \tag{4.13}$$

$$\sup_{\substack{t \geq u \geq 3 \\ y \in \mathbb{R}^2}} \widetilde{m}_n(y,u,t) \leq D_n, \tag{4.14}$$

and

$$\sup_{t \geq s \geq 3} \frac{1}{s} \int_{\mathbb{R}^2} dy\, \widetilde{m}_n(y,s,t) \leq E_n. \tag{4.15}$$

For $n = 1$ the assertions clearly hold because

$$\widetilde{m}_1(x_t, s_t, t) = t^{-\alpha} s_t \int_{B_t} p_{s_t}(x,y) dy \overset{t\to\infty}{\longrightarrow} \varphi(x)|B|, \tag{4.16}$$

$$\begin{aligned} \frac{1}{s_t}\int_{\mathbb{R}^2} \widetilde{m}_1(y, s_t, t)\, dy &= t^{-\alpha} \int_{B_t} dy \int_{\mathbb{R}^2} dz\, p_{s_t}(z,y) \\ &= t^{-\alpha} \int_{B_t} dy = |B|, \end{aligned} \tag{4.17}$$

$$\frac{1}{s}\int_{\mathbb{R}^2} (s-u) p_{s-u}(y,z) \widetilde{m}_1(z,u,t)\, dz = \frac{s-u}{s} u t^{-\alpha} \int_{B_t} p_s(y,z)\, dz \leq \frac{s-u}{s}\frac{u}{s}|B| \leq |B|, \tag{4.18}$$

$$\widetilde{m}_1(y,u,t) = ut^{-\alpha} \int_{B_t} p_u(y,z)\,dz \le |B|, \tag{4.19}$$

and

$$\frac{1}{s}\int_{\mathbb{R}^2} \widetilde{m}_1(y,s,t)\,dy = t^{-\alpha}\int_{\mathbb{R}^2} dy \int_{B_t} dz\, p_s(y,z) = |B|. \tag{4.20}$$

We will also need the following bound for the moments of the total mass

$$\mathbf{E}^x[(\eta_t(\mathbb{R}^2))^n] \le F_n \cdot (t+1)^{n-1}, \tag{4.21}$$

where $F_n = n!$. For $n=1$ this is clear since the l.h.s. of (4.21) equals 1. For $n \ge 2$ this is easily shown by induction using (4.11)

$$\begin{aligned} \mathbf{E}^x[(\eta_t(\mathbb{R}^2))^n] &\le F_1(t+1) + \frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k} F_k F_{n-k}\int_0^t (s+1)^{n-2}\,ds \qquad (4.22)\\ &= F_1(t+1) + \frac{1}{2}\frac{1}{n-1}\sum_{k=1}^{n-1}\binom{n}{k}F_kF_{n-k}(t+1)^{n-1}\\ &\le n!(t+1)^{n-1}. \end{aligned}$$

The uniformity of the claim in terms of the sequences (a_t), (b_t) and (c_t) will be needed to do the induction properly. Following the lines of the proof it can easily be established. We omit the details to avoid an unnecessary blow-up of the proof.

Let now $n \ge 2$. In the following we will assume that the validity of (4.7), (4.8) and (4.13)-(4.15) is already shown for all $n' < n$.

First note that

$$m_1(x_t,s_t,t) \ll \frac{\log t}{s_t}, \tag{4.23}$$

i.e. the l.h.s. in (4.23) is negligible compared with the expected main term. We thus calculate now

$$h_{n,k}(x,s,v,w) := \int_v^w du \int_{\mathbb{R}^2} dy\, p_{s-u}(x,y) m_k(y,u,t) m_{n-k}(y,u,t). \tag{4.24}$$

Let $(\delta_t)_{t\ge 0}$ be a sequence with $\delta_t \uparrow \infty$ so slowly that $\frac{\delta_t}{\log t} \stackrel{t\to\infty}{\longrightarrow} 0$. By (4.21)

$$\begin{aligned} h_{n,k}(x_t,s_t,0,\delta_t t^\alpha) &\le F_kF_{n-k}\int_0^{\delta_t t^\alpha}(u+1)^{n-2}\int_{\mathbb{R}^2} dy\, p_{s_t-u}(x_t,y)\int_{B_t} dz\, p_u(y,z) \qquad (4.25)\\ &\le \frac{F_kF_{n-k}}{n-1}(\delta_t t^\alpha+1)^{n-1}\frac{t^\alpha}{s_t}|B|\\ &\ll \frac{t^{n\alpha}}{s_t}(\log s_t)^{n-1} \qquad (4.26) \end{aligned}$$

is small. The other side of the integration interval will be estimated as follows. Let $(\varepsilon_t)_{t\ge0}$ be a sequence such that $\varepsilon_t \downarrow 0$ and such that $\frac{\log \varepsilon_t}{\log t} \stackrel{t\to\infty}{\longrightarrow} 0$. Then

$$\begin{aligned} h_{n,k}(x_t,s_t,\varepsilon_t s_t,s_t) &\le 2(C_k+D_k)D_{n-k}\frac{t^{n\alpha}}{s_t}\int_{\varepsilon_t s_t}^{s_t}\frac{(\log u)^{n-2}}{u}\,du \qquad (4.27)\\ &= 2(C_k+D_k)D_{n-k}\frac{1}{n-1}\frac{t^{n\alpha}}{s_t}(\log s_t)^{n-1}\left[1-\left(1-\frac{\log\varepsilon_t}{\log s_t}\right)^{n-1}\right]\\ &\ll \frac{t^{n\alpha}}{s_t}(\log s_t)^{n-1}. \end{aligned}$$

So the main term results from the integration over $[\delta_t t^\alpha, \varepsilon_t s_t]$. To evaluate this integral we split the spatial integral into the integral over the disc $D_u = \{y \in \mathbb{R}^2 : \|y\| \le K_u\sqrt{u}\}$ and its complement $D_u^c = \mathbb{R}^2 \setminus D_u$, where $K_u \uparrow \infty$ as $u \to \infty$ will be fixed later. By the induction hypotheses (4.7), (4.13) and (4.14) we get

$$\limsup_{t\to\infty} \frac{s_t t^{-n\alpha}}{(\log s_t)^{n-1}} \int_{\delta_t t^\alpha}^{\varepsilon_t s_t} du \, \frac{1}{u} \int_{D_u^c} dy \, p_{s_t-u}(x_t, y) \; u \; m_k(y,u,t) m_{n-k}(y,u,t) \tag{4.28}$$

$$\le D_{n-k} \limsup_{t\to\infty} \frac{1}{(\log s_t)^{n-1}} \int_{\delta_t t^\alpha}^{\varepsilon_t s_t} du \, \frac{(\log u)^{n-2}}{u} \int_{D_u^c} dy \, s_t \; p_{s_t-u}(x_t,y) \frac{1}{u} \widetilde{m}_k(y,u,t)$$

$$\le 2D_{n-k} \limsup_{t\to\infty} \frac{1}{(\log s_t)^{n-1}} \int_{\delta_t t^\alpha}^{\varepsilon_t s_t} du \, \frac{(\log u)^{n-2}}{u} \int_{D_u^c} dy \, \frac{1}{u} \widetilde{m}_k(y,u,t).$$

The last inequality holds since $s_t p_{s_t-u}(x_t, y) \le 2$ for $\varepsilon_t < \frac{1}{2}$. Fix $\beta' \ge 0$ and let (u_t) be a sequence such that $\frac{\log u_t}{\log t} \stackrel{t\to\infty}{\longrightarrow} \beta'$. Then by Fatou's lemma

$$\begin{aligned} \liminf_{t\to\infty} \int_{D_{u_t}} \frac{1}{u_t} \, \widetilde{m}_k(y,u_t,t)\, dy &= \liminf_{t\to\infty} \int_{\|y\|\le K_{u_t}} \widetilde{m}_k(y\sqrt{u_t}, u_t, t)\, dy \\ &\ge \left(1 - \frac{\alpha}{\beta'}\right)^{k-1} \frac{|B|^k k!}{(8\pi)^{k-1}} \cdot \int_{\mathbb{R}^2} \varphi(y)\, dy \\ &= \left(1 - \frac{\alpha}{\beta'}\right)^{k-1} \frac{|B|^k k!}{(8\pi)^{k-1}}. \end{aligned} \tag{4.29}$$

Let (u_t) be a sequence with $u_t \gg t^\alpha$ and let $\delta > 0$. Then by (4.8) for t sufficiently large

$$\int_{D_u^c} \frac{1}{u} \widetilde{m}_k(y,u,t)\, dy < \delta. \tag{4.30}$$

Thus the expression in (4.28) is less or equal than

$$\frac{2\delta D_{n-k}}{n-1} \limsup_{t\to\infty} \frac{(\log \varepsilon_t s_t)^{n-1} - (\log \delta_t t^\alpha)^{n-1}}{(\log s_t)^{n-1}} = \frac{2\delta D_{n-k}}{n-1} \left(1 - \left(\frac{\alpha}{\beta}\right)^{n-1}\right). \tag{4.31}$$

Since $\delta > 0$ was arbitrary the three expressions in (4.28) are equal and equal to zero.

Our task is now to determine the main term. By (4.7), (4.14) and the theorem of dominated convergence we may let $K_u \uparrow \infty$ so slowly that (uniformly in $\beta' \le 1$)

$$\begin{aligned} \frac{1}{u_t} \int_{D_u} \widetilde{m}_k(y,u_t,t) \widetilde{m}_{n-k}(y,u_t,t)\, dy &= \int_{\|y\|\le K_{u_t}} \widetilde{m}_k(y,u_t,t) \widetilde{m}_{n-k}(y,u_t,t)\, dy \\ &\stackrel{t\to\infty}{\longrightarrow} \left(1 - \frac{\alpha}{\beta'}\right)^{n-2} \int_{\mathbb{R}^2} \varphi(y)^2\, dy \cdot \frac{|B|^n k!(n-k)!}{(8\pi)^{n-2}} \\ &= 2\left(1 - \frac{\alpha}{\beta'}\right)^{n-2} \frac{|B|^n k!(n-k)!}{(8\pi)^{n-1}}. \end{aligned} \tag{4.32}$$

Assuming further $K_{\varepsilon_t s_t}\sqrt{\varepsilon_t} \stackrel{t\to\infty}{\longrightarrow} 0$ we get uniformly in $u \le \varepsilon_t s_t$ and $y \in D_u$

$$s_t p_{s_t-u}(x_t, y) \stackrel{t\to\infty}{\longrightarrow} \varphi(x). \tag{4.33}$$

We are now in the position to calculate

$$\lim_{t\to\infty} \frac{s_t t^{-n\alpha}}{(\log s_t)^{n-1}} \int\limits_{\delta_t t^\alpha}^{\varepsilon_t s_t} du \int\limits_{D_u} dy\, p_{s_t-u}(x_t, y) m_k(y,u,t) m_{n-k}(y,u,t) \tag{4.34}$$

$$= \lim_{t\to\infty} \frac{s_t}{(\log s_t)^{n-1}} \int\limits_{\delta_t t^\alpha}^{\varepsilon_t s_t} du\, \frac{(\log u)^{n-2}}{u} \int\limits_{D_u} dy\, p_{s_t-u}(x_t, y) \frac{1}{u} \widetilde{m}_k(y,u,t) \widetilde{m}_{n-k}(y,u,t)$$

$$= \lim_{t\to\infty} \frac{\varphi(x)}{(\log s_t)^{n-1}} \int\limits_{\delta_t t^\alpha}^{\varepsilon_t s_t} du\, \frac{(\log u)^{n-2}}{u} \int\limits_{D_u} dy\, \frac{1}{u} \widetilde{m}_k(y,u,t) \widetilde{m}_{n-k}(y,u,t) \tag{4.35}$$

$$= \varphi(x) 2 \frac{|B|^n k!(n-k)!}{(8\pi)^{n-1}} \lim_{t\to\infty} \frac{1}{(\log s_t)^{n-1}} \int\limits_{\delta_t t^\alpha}^{\varepsilon_t s_t} \frac{(\log u)^{n-2}}{u} \left(1 - \alpha \frac{\log t}{\log u}\right)^{n-2} du$$

$$= \varphi(x) \frac{2}{n-1} \frac{|B|^n k!(n-k)!}{(8\pi)^{n-1}} \lim_{t\to\infty} \frac{(\log(\varepsilon_t s_t) - \alpha \log t)^{n-1} - (\log(\delta_t t^\alpha) - \alpha \log t)^{n-1}}{(\log s_t)^{n-1}}$$

$$= \varphi(x) \frac{2}{n-1} \left(1 - \frac{\alpha}{\beta}\right)^{n-1} \frac{|B|^n k!(n-k)!}{(8\pi)^{n-1}}.$$

Summation over k in (4.12) now yields (4.7).

To show (4.8) we integrate (4.12)

$$\int\limits_{\mathbb{R}^2} m_n(x,s,t)\, dx = \int\limits_{\mathbb{R}^2} m_1(x,s,t)\, dx + \frac{1}{2} \sum_{k=1}^{n-1} \binom{n}{k} \int_0^s du \int\limits_{\mathbb{R}^2} dy\, m_k(y,u,t) m_{n-k}(y,u,t). \tag{4.36}$$

As above the first term is small and we have to evaluate

$$g_{n,k}(v,w) := \int_v^w du \int\limits_{\mathbb{R}^2} dy\, m_k(y,u,t) m_{n-k}(y,u,t). \tag{4.37}$$

For (δ_t) as above we get from (4.15) and (4.21) that

$$\begin{aligned} g_{n,k}(3, \delta_t t^\alpha) &\le t^{(n-k)\alpha} F_k \int_3^{\delta_t t^\alpha} du\, \frac{(\log u)^{n-k-1}}{u} (u+1)^{k-1} \int\limits_{\mathbb{R}^2} dy\, \widetilde{m}_{n-k}(y,u,t) \\ &\le \left(\frac{4}{3}\right)^{k-1} \frac{F_k E_{n-k}}{n-k} t^{(n-1)\alpha} (t\delta_t)^{k-1} (\log(\delta_t t^\alpha))^{n-k} \ll t^\alpha \end{aligned} \tag{4.38}$$

(note that $\frac{u+1}{u} \le \frac{4}{3}$ on the domain of integration). Let $\Delta = \mathrm{diam}(B)$. By assumption $\Delta < \infty$ which serves to show that

$$\begin{aligned}
g_{n,k}(0,3) &\le \int_0^3 du \int_{\mathbb{R}^2} dy\, m_n(y,u,t) \qquad (4.39)\\
&= \int_0^3 du \int_{[0,\Delta t^{\alpha/2}[^2} dy \sum_{l\in\mathbb{Z}^d} m_n(y+l\Delta t^{\alpha/2},u,t)\\
&\le \int_0^3 du \int_{[0,\Delta t^{\alpha/2}[^2} dy\, \mathbf{E}^y[(\eta_u(\mathbb{R}^2))^n]\\
&\le \Delta^2 t^\alpha \frac{4^n-1}{n} F_n.
\end{aligned}$$

Since the expected main term is of order $t^{n\alpha}(\log t)^{n-1}$ we have got that $g_{n,k}(0,\delta_t t^\alpha)$ is negligible. Let also (ε_t) be as above to obtain by (4.14) and (4.15) that

$$\begin{aligned}
g_{n,k}(\varepsilon_t s_t, s_t) &= t^{n\alpha} \int_{\varepsilon_t s_t}^{s_t} du\, \frac{(\log u)^{n-2}}{u} \int_{\mathbb{R}^2} dy\, \widetilde{m}_k(y,u,t) \frac{1}{u} \widetilde{m}_{n-k}(y,u,t) \qquad (4.40)\\
&\le \frac{D_k E_{n-k}}{n-1} t^{n\alpha} \left[(\log s_t)^{n-1} - (\log(\varepsilon_t s_t))^{n-1}\right] \ll t^{n\alpha}(\log s_t)^{n-1}
\end{aligned}$$

is small. We split up $g_{n,k}(\delta_t t^\alpha, \varepsilon_t s_t)$ as above. The part resulting from the integral over D_u^c is small since

$$\begin{aligned}
&\int_{\varepsilon_t t^\alpha}^{\varepsilon^t s_t} du \int_{D_u^c} dy\, m_k(y,u,t) m_{n-k}(y,u,t) \qquad (4.41)\\
&\qquad \le D_{n-k} t^{n\alpha} \int_{\delta_t t^\alpha}^{\varepsilon_t s_t} du\, \frac{(\log u)^{n-2}}{u} \int_{D_u^c} dy\, \frac{1}{u} \widetilde{m}_k(y,u,t)\\
&\qquad \ll t^{n\alpha}(\log s_t)^{n-1}.
\end{aligned}$$

The integral over D_u has already been determined in (4.35).

So far we have shown part (a) of the lemma. To prove part (b) we still have to show that (4.13)-(4.15) hold and that the size of the constants can be controlled. We will do this by means of recursion formulas for C_n, D_n and E_n. Note that

$$m_k(x,u,t) \le F_k (u+1)^{k-1} \int_{B_t} p_u(x,y) dy.$$

Therefor we have

$$\begin{aligned}
&\int_0^3 du \int_{\mathbb{R}^2} dy\, (s-u) p_{s-u}(y,z) m_k(z,u,t) m_{n-k}(z,u,t) \qquad (4.42)\\
&\qquad \le F_k F_{n-k} \int_0^3 du\, (u+1)^{n-2} \int_{\mathbb{R}^2} dz \int_{B_t} dw\, (s-u) p_{s-u}(y,z) p_u(z,w)\\
&\qquad \le F_k F_{n-k} \int_0^3 du\, (u+1)^{n-2} \int_{B_t} dw\, (s-u) p_s(z,w)\\
&\qquad \le \frac{F_k F_{n-k}}{n-1} 4^{n-1} |B| t^\alpha.
\end{aligned}$$

Putting this into the recursion formula (4.12) we get

$$\frac{s-u}{u}\int_{\mathbb{R}^2} p_{s-u}(y,z)\widetilde{m}_n(z,u,t)\,dz \le \frac{1}{(\log u)^{n-1}}\left[C_1+\frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\left(\frac{F_kF_{n-k}}{n-1}4^{n-1}|B|\right.\right. \tag{4.43}$$
$$\left.\left.+\int_{\mathbb{R}^2} dz\int_3^u dv\,\frac{(\log v)^{n-2}}{v}\int_{\mathbb{R}^2} dz'\,(s-u)p_{s-u}(y,z)p_{u-v}(z,z')\frac{1}{v}\widetilde{m}_k(z',v,t)\widetilde{m}_{n-k}(z',v,t)\right)\right].$$

Doing the integration the summands equal

$$\begin{aligned}\int_3^u dv\,\frac{(\log v)^{n-2}}{v}\int_{\mathbb{R}^2} dz'\,(s-u)p_{s-v}(y,z')\frac{1}{v}\widetilde{m}_k(z',v,t)\widetilde{m}_{n-k}(z',v,t) \\ \le C_kD_{n-k}\int_3^u\frac{(\log v)^{n-2}}{v}\,dv \le \frac{C_kD_{n-k}}{n-1}(\log u)^{n-1}.\end{aligned} \tag{4.44}$$

We have shown that (4.13) holds with

$$C_n \le C_1+\frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\left(\frac{C_kD_{n-k}}{n-1}+\frac{F_kF_{n-k}}{n-1}4^{n-1}|B|\right). \tag{4.45}$$

We now turn to the D_n. By the recursion formula (4.12) we get for $t\ge s\ge 3$ and $y\in\mathbb{R}^2$

$$\widetilde{m}_n(y,s,t)\le t^{-n\alpha}\frac{s}{(\log s)^{n-1}}\left(m_1(y,s,t)+\frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}h_{n,k}(y,s,0,t)\right). \tag{4.46}$$

Now

$$\begin{aligned} h_{n,k}(y,s,3,t) &\le D_{n-k}\int_0^2 du\,\frac{1}{u}\int_{real^2} dz\,p_{u-s}(y,z)m_k(z,s,t)(\log su)^{n-k-1} \\ &\le 2(C_k+D_k)D_{n-k}\int_3^s\frac{(\log u)^{n-2}}{u}du \\ &\le \frac{2(C_k+D_k)D_{n-k}}{n-1}\frac{(\log s)^{n-1}}{s}.\end{aligned} \tag{4.47}$$

From this and (4.25) we get that D_n can be chosen to be

$$D_n\le D_1+\frac{1}{2(n-1)}\sum_{k=1}^{n-1}\binom{n}{k}\left[F_kF_{n-k}4^{n-1}|B|+2(C_k+D_k)D_{n-k}\right]. \tag{4.48}$$

Finally the E_n will be determined as follows

$$\frac{1}{s}\int_{\mathbb{R}^2}\widetilde{m}_n(y,s,t)\,dy\le E_1+\frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\frac{1}{(\log s)^{n-1}}g_{n,k}(0,s). \tag{4.49}$$

Now

$$\begin{aligned} g_{n,k}(3,s) &\le D_k\int_3^s\frac{1}{u}(\log u)^{k-1}\int_{\mathbb{R}^2} dz\,m_{n-k}(z,u,t) \\ &\le D_kE_{n-k}\int_3^s\frac{(\log u)^{k-2}}{u}du \\ &\le \frac{D_kE_{n-k}}{n-1}(\log s)^{n-1}.\end{aligned} \tag{4.50}$$

Together with (4.39) this yields that we can choose E_n to be

$$E_n \leq E_1 + \frac{1}{2(n-1)} \sum_{k=1}^{n-1} \binom{n}{k} \left[D_k E_{n-k} + \Delta^2 4^n F_n \right]. \tag{4.51}$$

Putting together (4.21),(4.45), (4.48) and (4.51) we see that we can choose

$$C_n = D_n = E_n = n! \Gamma^n \tag{4.52}$$

for some $\Gamma < \infty$ (depending on Δ). □

Since we will need some uniformity in different spatial scalings that are $\approx t^{\alpha/2}$ (recall that $a_t \approx b_t$ means $(\log a_t)/(\log b_t) \stackrel{t\to\infty}{\longrightarrow} 1$) we state one more lemma.

Lemma 4.2 *Let (ψ_t) be BBM($\mathbb{R}^2$) or SBM($\mathbb{R}^2$) and $I = [0,1]$ resp. $]-\infty, 1]$. Fix $\alpha \in I$ and $v(t) \ll u(t)$ with $u(t), v(t) \approx t^\alpha$. Then uniformly in all sequences $w(t)$ such that $u(t) \leq w(t) \leq v(t) \, \forall \, t \geq 0$ the following holds*

$$h(t) := \mathbf{E}^{\widetilde{M}(t)} \left[\left(\frac{1}{u(t)} \widetilde{\psi}_t([0, \sqrt{u(t)}[^2) - \frac{1}{w(t)} \widetilde{\psi}_t([0, \sqrt{w(t)}[^2) \right)^2 \right] \stackrel{t\to\infty}{\longrightarrow} 0. \tag{4.53}$$

Proof Let

$$\phi_t = \frac{1}{u(t)} 1\!\!1_{[0, u(t)^{1/2}[^2} - \frac{1}{w(t)} 1\!\!1_{[0, w(t)^{1/2}[^2}.$$

Then by the second moment formulas (3.5) and (3.12) (recall that (S_s) is the semigroup of Brownian motion on $\mathbb{R}^2$)

$$h(t) \leq a_t + b_t + c_t, \tag{4.54}$$

(with equality in the case of BBM) where

$$\begin{aligned} a_t &= \left(\frac{8\pi}{\log t} \right)^2 \int \left(\langle \mu, S_t \phi_t \rangle \right)^2 \widetilde{M}(t)(d\mu) \\ b_t &= \left(\frac{8\pi}{\log t} \right)^2 \int \left\langle \mu, S_t(\phi_t^2) - (S_t \phi_t)^2 \right\rangle \widetilde{M}(t)(d\mu) \\ c_t &= \left(\frac{8\pi}{\log t} \right)^2 \int \left\langle \mu, \int_0^T S_{t-s}((S_s \phi_t)^2) \, ds \right\rangle \widetilde{M}(t)(d\mu). \end{aligned} \tag{4.55}$$

Clearly $a_t \stackrel{t\to\infty}{\longrightarrow} 0$, $b_t \stackrel{t\to\infty}{\longrightarrow} 0$. For $c_t \stackrel{t\to\infty}{\longrightarrow} 0$ we have to be more careful. By translation inavariance we get (recall that λ is the Lebesgue measure)

$$c_t = \frac{8\pi}{\log t} \left\langle \lambda, \int_0^t (S_s \phi_t)^2 \, ds \right\rangle. \tag{4.56}$$

Note that by Hölder's inequality

$$\begin{aligned} \langle \lambda, (S_s \phi_t)^2 \rangle &\leq \| S_s \phi_t \|_\infty = \sup_{x \in \mathbb{R}^2} |S_s \phi_t(x)| \\ &\leq \min \left(\frac{1}{2\pi s}, \frac{1}{u(t)} + \frac{1}{w(t)} \right) \leq \min \left(\frac{1}{2\pi s}, \frac{2}{u(t)} \right). \end{aligned} \tag{4.57}$$

Thus

$$\frac{8\pi}{\log t}\int_0^{v(t)\log t} \langle \lambda, (S_s\phi_t)^2\rangle\, ds \quad \leq \quad \frac{8\pi}{\log t}\left[\frac{2}{\log t} + (\log(v(t)\log t) - \log(u(t)/\log t))\right] \quad \stackrel{t\to\infty}{\longrightarrow} \quad 0. \tag{4.58}$$

On the other hand

$$\begin{aligned} \|S_s\phi_t\|_\infty &\leq \sup_{x\in\mathbb{R}^2}\ \sup_{y\in[0,u(t)^{1/2}[^2}\ \sup_{z\in[0,v(t)^{1/2}[^2} |p_s(x,y)-p_s(x,z)| \qquad (4.59)\\ &= \frac{1}{2\pi s}\sup_{r\in\mathbb{R}}\ \sup_{\zeta\in[-(2v(t))^{1/2},(2v(t))^{1/2}]} \left|\exp\{-r^2/2s\} - \exp\{-(r-\zeta)^2/2s\}\right| \\ &\leq \frac{e^{-1}}{2\pi s}\sqrt{2v(t)/s}\,. \end{aligned}$$

Thus

$$\frac{8\pi}{\log t}\int_{v(t)\log t}^{t} \langle \lambda (S_s\phi_t)^2\rangle\, ds \leq \frac{\sqrt{8}}{e}\sqrt{1/\log t} \stackrel{t\to\infty}{\longrightarrow} 0. \tag{4.60}$$

We conclude $c_t \stackrel{t\to\infty}{\longrightarrow} 0$ and the proof is complete. □

5 Proof of the Clustering Results for the Infinite Systems

5.1 Proof of Theorem 1

The proof of Theorem 1 will be based on an asymptotic result related to the Laplace transform. This is formulated in Proposition 5.1 and 5.2 below.

Let $x \in \mathbb{R}^2$ and $(x_t)_{t\geq 0}$ a sequence in $\mathbb{R}^2$ such that $\frac{x_t}{\sqrt{t}} \stackrel{t\to\infty}{\longrightarrow} x$.

Proposition 5.1 *Then for $B \in \mathcal{B}(\mathbb{R}^2)$ and $\theta \geq 0$*

$$\lim_{t\to\infty}\frac{t\log t}{8\pi}\left(1-\mathbf{E}^{x_t}\left[\exp\{-\theta\widetilde{\psi}_t^\alpha(B)\}\right]\right) \stackrel{t\to\infty}{\longrightarrow} \varphi(x)\frac{\theta|B|}{1+\theta|B|(1-\alpha)} \tag{5.1}$$

$$\lim_{t\to\infty}\frac{\log t}{8\pi}\left(1-\mathbf{E}^{M(1)}\left[\exp\{-\theta\widetilde{\psi}_t^\alpha(B)\}\right]\right) \stackrel{t\to\infty}{\longrightarrow} \frac{\theta|B|}{1+\theta|B|(1-\alpha)}. \tag{5.2}$$

Proposition 5.1 can be reformulated in terms of distributions.

Proposition 5.2 *Let (x_t) as in Proposition 5.1 and let $u > 0$. Then for $B \in \mathcal{B}(\mathbb{R}^2)$ and $\theta \geq 0$*

$$\lim_{t\to\infty}\frac{t\log t}{8\pi}\mathbf{P}^{x_t}\left[\widetilde{\psi}_t^\alpha(B) > u\right] = \frac{\varphi(x)}{1-\alpha}\exp\left\{-\frac{u}{|B|(1-\alpha)}\right\} \tag{5.3}$$

$$\lim_{t\to\infty}\frac{\log t}{8\pi}\mathbf{P}^{M(1)}\left[\widetilde{\psi}_t^\alpha(B) > u\right] = \frac{1}{1-\alpha}\exp\left\{-\frac{u}{|B|(1-\alpha)}\right\}. \tag{5.4}$$

Proof (of Proposition 5.1)

Let

$$\phi_t(\theta) = \frac{t\log t}{8\pi}\Big(1 - \mathbf{E}^{x_t}[\exp\{-\theta\widetilde{\psi}_t^\alpha(B)\}]\Big) \qquad \theta \in \mathbb{C},\ \mathrm{Re}(\theta) > 0. \tag{5.5}$$

Then

$$|\phi_t(\theta)| \le \frac{t\log t}{8\pi}|\theta| \cdot \mathbf{E}^{x_t}[\widetilde{\psi}_t^\alpha(B)] \le |\theta|. \tag{5.6}$$

Thus $\phi_t(\theta)$ is uniformly bounded for θ in compact sets. Let $\Gamma < \infty$ be as in Lemma 4.1(b). By (4.9) for $|\theta| < \frac{1}{\Gamma}$ we can express $\phi_t(\theta)$ in terms of the moments

$$\begin{aligned} \phi_t(\theta) &= \frac{t\log t}{8\pi}\sum_{n=1}^{\infty}\frac{(-\theta)^n\mathbf{E}^{x_t}[(\widetilde{\psi}_t^\alpha(B))^n]}{n!} \\ &= \sum_{n=1}^{\infty}\frac{(-\theta)^n(8\pi)^{n-1}\widetilde{m}_n(x_t,s_t,t,\alpha)}{n!}. \end{aligned} \tag{5.7}$$

Hence by (4.7)

$$\phi_t(\theta) \overset{t\to\infty}{\longrightarrow} \frac{\theta|B|}{1+\theta|B|(1-\alpha)} \qquad ,\ |\theta| < \frac{1}{\Gamma}. \tag{5.8}$$

By Vitali's theorem (see e.g. Remmert (1991)) equation (5.8) holds for all θ on the right half plane.

The proof of (5.2) is analogous. Here we take

$$\phi_t(\theta) = \frac{\log t}{8\pi}\left[1 - \mathbf{E}^{M(1)}\left[\exp\{-\theta\widetilde{\psi}_t^\alpha(B)\}\right]\right] \tag{5.9}$$

and use (4.8) and (4.10). □

Proof (of Theorem 1)

From Proposition 5.1 the proof is easy. From (1.7) the Laplace transform

$$L(s,\theta) = \mathbf{E}^1[\exp\{-\theta Z_s\}] \tag{5.10}$$

of the Feller diffusion (Z_s) solves

$$\begin{aligned} \frac{\partial}{\partial s}L(s,\theta) &= \mathbf{E}^1[\theta^2 Z_s\exp\{-\theta Z_s\}] = -\theta^2\frac{\partial}{\partial\theta}L(s,\theta) \\ L(0,\theta) &= \exp\{-\theta\}. \end{aligned} \tag{5.11}$$

The solution of (5.11) is

$$L(s,\theta) = \exp\left\{-\frac{\theta}{1+\theta s}\right\}, \qquad \theta \ge 0,\ s \ge 0. \tag{5.12}$$

Let $\alpha \in [0,1]$. Use (5.2) to obtain

$$\begin{aligned} \mathbf{E}^{\widetilde{M}(t)}\left[\exp\{-\theta\widetilde{\psi}_t^\alpha(B)\}\right] &= \left(1 - \left(1 - \mathbf{E}^{M(1)}\left[\exp\{-\theta\widetilde{\psi}_t^\alpha(B)\}\right]\right)\right)^{\frac{\log t}{8\pi}} \\ &\overset{t\to\infty}{\longrightarrow} \exp\left\{-\frac{\theta|B|}{1+\theta|B|(1-\alpha)}\right\}. \end{aligned} \tag{5.13}$$

Comparing this with (5.12) yields the claim.

The case $\alpha < 0$ and $\psi_t = \zeta_t$ SBM($\mathbb{R}^d$) can be done with the scaling property (1.13) as follows

$$\begin{aligned}
\mathcal{L}^{\widetilde{M}(t)}\left[\widetilde{\zeta}_t^\alpha(B)\right] &= \mathcal{L}^{\widetilde{M}(t)}\left[\frac{8\pi}{\log t}t^{-\alpha}\zeta_t(t^{\alpha/2}B)\right] \\
&= \mathcal{L}^{\widetilde{M}(t)}\left[\frac{8\pi}{\log t}\zeta_{t^{1-\alpha}}(B)\right] \\
&= \mathcal{L}^{\widetilde{M}(t^{1-\alpha})/(1-\alpha)}\left[(1-\alpha)\widetilde{\zeta}_{t^{1-\alpha}}(B)\right] \qquad (5.14)\\
&\overset{t\to\infty}{\Longrightarrow} \mathcal{L}^{1/(1-\alpha)}\left[(1-\alpha)Z_1\right] = \mathcal{L}^1[Z_{1-\alpha}].
\end{aligned}$$

□

5.2 Proof of Theorem 2

It is sufficient to check that

$$\mathcal{L}^{\widetilde{M}(t)}\left[(\mathcal{S}_{A(e),t}\mathcal{T}_{x_t^e}\widetilde{\psi}_t(B^e))_{e\in\mathbb{T}}\right] \overset{t\to\infty}{\longrightarrow} \mathcal{L}\left[\left(|B_e|Z^e_{1-A(e)}\right)_{e\in\mathbb{T}}\right], \qquad (5.15)$$

for $B^e \in \mathcal{B}(\mathbb{R}^2)$ bounded for all $e \in \mathbb{T}$.

We do the proof by induction over the length of the tree $\mathbb{T}$. For $\mathbb{T} = \{\emptyset\}$ this is the assertion of Theorem 1 (together with Lemma 4.2). Now assume that the claim has been shown for all trees shorter than $\mathbb{T}$.

The idea of the proof is the following. We introduce a time scale $L(t) \approx t^{A(\emptyset)}$ and couple (ψ_s) for $s \geq t - L(t)$ with another process (ψ_s^2). This process shall have initial configuration $M(\rho)$, where ρ is the empirical population density of $\psi^1_{t-L(t)}$ in a box of length $\approx t^{A(\emptyset)/2}$. $L(t)$ will be chosen small enough that the evolutions of the subtrees (resulting from eliminating $\emptyset$ from $\mathbb{T}$) are approximately independent. On the other hand $L(t)$ has to be chosen large enough so that the local coupling with local size $R(t) \approx t^{A(\emptyset/2)}$ is successful. Here a the details.

Let $b = \max\{\mathrm{diam}(B^e),\, e \in \mathbb{T}\}$. Let $d_t \downarrow 0$, $t \to \infty$ such that

$$\begin{aligned}
t^{(A(e\wedge f)-d_t)/2} &\leq \|x_t^e - x_t^f\| - b(t^{A(e)/2} + t^{A(f)/2}) \qquad (5.16)\\
&\leq \|x_t^e - x_t^f\| + b(t^{A(e)/2} + t^{A(f)/2}) \leq \frac{1}{2}t^{(A(e\wedge f)+d_t)/2}
\end{aligned}$$

for all $e, f \in \mathbb{T}$. We may and will assume that $t^{d_t} \overset{t\to\infty}{\longrightarrow} \infty$. Let $\alpha := A(\emptyset)$. Let

$$\begin{aligned}
S = S(t) &= t^{(\alpha+d_t)/2}\\
R = R(t) &= t^{(\alpha-3d_t)/2}\\
L = L(t) &= t^{\alpha-2d_t}.
\end{aligned}$$

Let

$$B_t^e = x_t^e + t^{A(e)/2}B^e \qquad (5.17)$$

and

$$B_t = \bigcup_{e\in\mathbb{T}} B_t^e. \qquad (5.18)$$

By shifting $X = (x^e_t,\, e \in \mathbb{T})$ if necessary we can assume that $B_t \subset [0,S[^2$ for all $t > 0$ and

$$L^{-1/2} \cdot \mathrm{dist}(B_t, \mathbb{R}^2 \setminus [0,S[^2[) \stackrel{t\to\infty}{\longrightarrow} \infty. \tag{5.19}$$

Apply Corollary 3.7 with $\psi^1_0 = \psi_{t-L(t)}$, $s = L(t)$, $\rho = \log t/8\pi$ and with $\varepsilon = \delta = \frac{\log t}{8\pi}\varepsilon_t$, where $\varepsilon_t \stackrel{t\to\infty}{\longrightarrow} 0$. This last choice is possible due to Lemma 4.2. Thus we obtain a coupling $(\psi^1_s, \psi^2_s)_{s\geq 0}$ with $\mathcal{L}[\psi^1_0|\psi^1_s] = M(S^{-2}\psi^1_0([0,S[^2))$ such that there exists a sequence $\delta_t \downarrow 0$ with

$$\mathbf{E}^{\widetilde{M}(t)}\left[\left|(\widetilde{\psi}^1_{L(t)} - \widetilde{\psi}^2_{L(t)})(C)\right|\right] \leq \delta_t \cdot |C| \quad \forall\, C \in \mathcal{B}(\mathbb{R}^2). \tag{5.20}$$

So all we have to show is

$$\mathcal{L}^{\widetilde{M}(t)}\left[\frac{8\pi}{\log t}\left(t^{-A(e)}\psi^2_{L(t)}(B^e_t)\right)_{e\in\mathbb{T}}\right] \stackrel{t\to\infty}{\Longrightarrow} \mathcal{L}^1\left[(|B^e|Z^e_{1-A(e)})_{e\in\mathbb{T}}\right]. \tag{5.21}$$

By Theorem 1 (and Lemma 4.2) we know that

$$\mathcal{L}^{\widetilde{M}(t)}\left[\frac{8\pi}{\log t}S^{-2}\psi^1_0([0,S[^2)\right] \stackrel{t\to\infty}{\Longrightarrow} \mathcal{L}[Z_{1-\alpha}]. \tag{5.22}$$

Hence (using the Chapman-Kolmogorov equation) showing (5.21) amounts to showing for $\rho \geq 0$

$$\begin{aligned} \mathcal{L}^{M(\rho\log t/8\pi)}\left[\frac{8\pi}{\log t}\left(t^{-A(e)}\psi_{L(t)}(B^e_t)\right)_{e\in\mathbb{T}}\right] &\stackrel{t\to\infty}{\Longrightarrow} \mathcal{L}^\rho\left[(Z^e_{\alpha-A(e)})_{e\in\mathbb{T}}\right] \\ &= \mathcal{L}^{\rho/\alpha}\left[(\alpha Z^e_{1-A(e)/\alpha})_{e\in\mathbb{T}}\right]. \end{aligned} \tag{5.23}$$

The last equality is the basic scaling property of Feller's diffusion.

Let $\mathbb{T}_j = \{(j, l_2, \ldots, l_n) \in \mathbb{T}, n \in \mathbb{N}\}$, $j = 1, \ldots, J$ be the partition of $\mathbb{T}$ into subtrees $\mathbb{T}_j$ ($\mathbb{T} = \{\emptyset\} \cup \mathbb{T}_1 \cup \cdots \cup \mathbb{T}_J$). To prove (5.23) it suffices (by the induction hypothesis) to show that

$$\left(\tfrac{8\pi}{\log t}t^{-A(e)}\psi_{L(t)}(B^e_t)\right)_{e\in\mathbb{T}_j}, \quad j = 1, \ldots, J \tag{5.24}$$

are J asymptotically independent random variables.

Fix one $e_j \in \mathbb{T}_j$ for each $j = 1, \ldots, J$ and let $C_j = C_j(t) = x^{e_j}_t + [-R(t), R(t)[^2$ and $C_0 = \mathbb{R}^2 \setminus (C_1 \cup \cdots \cup C_J)$. Then for t large enough we have $C_i \cap C_j = \emptyset$ for $i \neq j$. Let

$$\Delta_j = \Delta_j(t) = \inf_{e\in\mathbb{T}_j} \mathrm{dist}(B^e_t, \mathbb{R}^2 \setminus C_j).$$

Since $A : \mathbb{T} \to I$ is strictly decreasing we have $\Delta_j(t)/\sqrt{L(t)} \stackrel{t\to\infty}{\longrightarrow} \infty$.

Let $(\chi^j_s)_{s\geq 0}$, $j = 0, 1, \ldots, J$ be independent BBM($\mathbb{R}^2$) resp. SBM($\mathbb{R}^2$) with $\chi^j_0 = M(\frac{\log t}{8\pi}\rho) \cdot 1\!\!1_{C_j}$, $j = 0, 1, \ldots, J$. We can assume

$$\psi_s = \chi^0_s + \cdots + \chi^J_s.$$

Now for $j = 1, \ldots, J$ and $e \in \mathbb{T}_j$

$$\begin{aligned} &\mathbf{E}\left[\frac{8\pi}{\log t}t^{-A(e)}\sum_{\substack{i=0\\ i\neq j}}^{J}\chi^i_{L(t)}(B^e_t)\right] \qquad (5.25) \\ &\quad\leq \rho|B^e|t^{-A(e)}\int\limits_{\mathbb{R}^2\setminus C_j} dx \int\limits_{B^e_t} dy\, p_{L(t)}(x,y) \leq \rho|B^e|\exp\{-\Delta^2_j/L(t)\} \stackrel{t\to\infty}{\longrightarrow} 0. \end{aligned}$$

Thus (5.24) holds and the proof is complete. □

6 Proofs for Finite Systems

6.1 Proof of Theorem 3

The idea of the proof is again to introduce a new time scale $L(t) \ll t^2$ (recall Λ_t^d has width t) and to let $T'(t) = T(t) - L(t)$. As in the previous section we want to couple (locally) given $t^{-d}\psi_{T'(t)}(\Lambda_t^d) = \rho$ with a process started in $M_t(\rho)$. This latter one will then be compared to the infinite process started in $M(\rho)$. So as to impose the local coupling we will have to cut Λ_t^d into a growing (with t) number $N(t)^d$ of boxes. $N(t)$ has to be chosen such that the empirical densities of $\psi_{T'(t)}$ within the boxes and within Λ_t^d are asymptotically close.

Step 1. We start with showing this latter point. Let $A, B \in \mathcal{B}(\Lambda_1^d)$, $|A|, |B| > 0$ and $\phi_t = \frac{1}{|tA|}1\!\!1_{tA} - \frac{1}{|tB|}1\!\!1_{tB}$ for $t > 0$. Then by the second moment formulas (3.5) and (3.12) (recall that (S_s) is the semigroup and $p_{t,s}(\cdot,\cdot)$ the transition density of Brownian motion on Λ_t^d)

$$\mathbf{E}^{M_t(\rho)}\left[\left(\frac{1}{|tA|}\psi_{t,T(t)}(tA) - \frac{1}{|tB|}\psi_{t,T(t)}(tB)\right)^2\right] \tag{6.1}$$
$$\leq \int \left(\langle \mu, S_{T(t)}\phi_t\rangle\right)^2 + \left\langle \mu, S_{T(t)}(\phi_t^2) - (S_{T(t)}\phi_t)^2\right\rangle + \left\langle \mu, \int_0^T S_{T(t)-s}(S_s\phi_t)^2\,ds\right\rangle M_t(\rho)(d\mu)$$

with equality in the case of BBM. Fix a sequence $\gamma(t)$ such that $t^2 \ll \gamma(t) \ll T(t)$. Then

$$\sup_{u\geq\gamma(t)}\sup_{z\in\Lambda_t^d}|t^d p_{t,u}(0,z) - 1| =: \varepsilon_t \stackrel{t\to\infty}{\longrightarrow} 0. \tag{6.2}$$

Thus for $u \geq \gamma(t)$

$$\sup_{x\in\Lambda_t^d}|\langle\delta_x, S_u\phi_t\rangle| \leq 2\varepsilon_t t^{-d} \tag{6.3}$$

and, of course, for all $u \geq 0$

$$\sup_{x\in\Lambda_t^d}|\langle\delta_x, S_u\phi_t\rangle| \leq \left(\frac{1}{|A|} + \frac{1}{|B|}\right)t^{-d}. \tag{6.4}$$

Note that $\phi_t^2 \leq t^{-2d}(\frac{1}{|A|} + \frac{1}{|B|})^2$. Hence (6.1) is dominated by

$$4\varepsilon_t^2 t^{-2d}(\rho^2 t^{2d} + \rho t^d) + \rho\left(\frac{1}{|A|} + \frac{1}{|B|}\right)^2 t^{-d} + \rho\left[\varepsilon_t^2 T(t)t^{-d} + \left(\frac{1}{|A|} + \frac{1}{|B|}\right)^2\gamma(t)t^{-d}\right] \stackrel{t\to\infty}{\longrightarrow} 0. \tag{6.5}$$

If we replace $T(t)$ by $T'(t)$ this convergence is uniform in all sequences $T'(t)$ such that $\frac{1}{2}T(t) \leq T'(t) \leq T(t)$. Thus we can find a sequence $N(t) \uparrow \infty$, $\frac{\log N(t)}{\log t} \stackrel{t\to\infty}{\longrightarrow} 0$ and define $L(t) = \frac{t^2}{N(t)}$, $T'(t) = T(t) - L(t)$ such that

$$t^{-2d}\mathbf{E}^{M_t(\rho)}\left[\left|\psi_{t,T'(t)}(\Lambda_t^d) - N(t)^d\psi_{t,T'(t)}\left(\left[0, \tfrac{t}{N(t)}\right[^d\right)\right|\right] =: \delta_t \stackrel{t\to\infty}{\longrightarrow} 0. \tag{6.6}$$

Step 2. *(Coupling)* We continue arguing as in the proof of Theorem 2. We let $(\chi_{t,s}^1, \chi_{t,s}^2)_{s\geq 0}$ be the local coupling of BBM(Λ_t^d) resp. SBM(Λ_t^d) according to Corollary 3.6 with $R = R(t) = \frac{t}{N(t)}$. The initial configuration shall be $\chi_{t,0}^1 = \psi_{t,T'(t)}$ and

$\mathcal{L}[\chi_0^2|\chi_0^1] = M_t(t^{-2}\chi_0^1(\Lambda_t^d))$. By Corollary 3.6 we get for $B \in \mathcal{B}(\mathbb{R}^d)$ bounded

$$\mathbf{E}^{M_t(\rho)}\left[\left\|(\chi^1_{t,L(t)} - \chi^2_{t,L(t)})\Big|_B\right\|\right] \le |B| \cdot \left[\delta_t + 2\sqrt{\rho R(t)^{-d}} + 2\sqrt{\frac{d}{\pi}}\,\rho N(t)^{-1/2}\right] \overset{t\to\infty}{\longrightarrow} 0. \quad (6.7)$$

Step 3. (Comparison) We apply the comparison lemma (Lemma 3.8) to $(\chi_s^3)_{s\ge 0}$ with $\mathcal{L}[\chi_0^3|\chi_0^1] = M(t^{-d}(\Lambda_t^d))$ and $(\chi^2_{t,s})$ and with $A_t \equiv B$ to obtain

$$\mathbf{E}\left[\left|\chi^2_{t,L(t)}(B) - \chi^3_{L(t)}(B)\right|\right] \overset{t\to\infty}{\longrightarrow} 0. \quad (6.8)$$

Thus

$$\mathbf{E}\left[\left|\chi^1_{t,L(t)}(B) - \chi^3_{L(t)}(B)\right|\right] \overset{t\to\infty}{\longrightarrow} 0. \quad (6.9)$$

Step 4. (Conclusion) Fix $f \in C_c(\mathbb{R}^d)$ and $F \in C_b(\mathbb{R})$. Then

$$\begin{aligned}\mathbf{E}^{M_t(\rho)}[F(\langle \psi_{t,T(t)}, f\rangle)] &= \mathbf{E}[F(\langle \chi^1_{t,L(t)}, f\rangle)] \quad (6.10)\\ &= \mathbf{E}[F(\langle \chi^2_{t,L(t)}, f\rangle)] + o(1)\\ &= \mathbf{E}[F(\langle \chi^3_{t,L(t)}, f\rangle)] + o(1)\\ &= \int_0^\infty \mathbf{P}^\rho[Z_{\sigma/2} \in d\rho']F(\langle \nu_{\rho'}, f\rangle) + o(1).\end{aligned}$$

The last equality holds because of (1.16) and (2.9). □

6.2 Proof of Theorem 4 and 5

The proofs are similar to that of Theorem 3. Hence we give only an outline. Recall $\beta(t) = t^2 \log t$. By (2.9) we know that

$$\mathcal{L}^{\widetilde{M}_t(\beta(t))}\left[\frac{8\pi}{\log \beta(t)} t^{-2} \|\psi_{t,T'(t)}\|\right] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}^1[Z_{2\pi\sigma}]. \quad (6.11)$$

Choose $L(t) \ll t^2$ such that $\lim_{t\to\infty} \frac{\log L(t)}{\log \beta(t)} = \lim_{t\to\infty} \frac{\log L(t)}{\log t^2} = 1$. Now we can proceed as in the proof of Theorem 3. We couple locally with the configuration

$$\int_0^\infty \mathbf{P}^1[Z_{2\pi\sigma} \in d\rho] M_t\left(\rho\,\frac{\log \beta(t)}{8\pi}\right) \quad (6.12)$$

and compare this with the infinite system started in

$$\int_0^\infty \mathbf{P}^1[Z_{2\pi\sigma} \in d\rho] M\left(\rho\,\frac{\log \beta(t)}{8\pi}\right). \quad (6.13)$$

Now we apply Theorem 1 resp. 2 to obtain the conclusions. □

Appendix

1 Description of the Simulations

We give a short description how the simulations of the Figures I.1, I.2 and III.3 have been generated.

Voter Model, Direct Approach

The direct approach to simulate a voter model on the finite site space S is the following: Start with each site $x \in S$ being randomly coloured white or black. Repeat the following procedure

- Pick one point $x \in S$ at random.
- Choose a "neighbour" y at random according to the interaction kernel $a(\cdot, \cdot)$.
- Change x's opinion to that of y.

This method has been applied to obtain the voter model in Figure I.2 where we wanted to see the evolution in time. However, proceeding like this has the disadvantage of being hopelessly slow.

Voter Model via Duality

For the long-time simulation of Figure I.1 we needed something faster. Here we exploited the fact that Figure I.1 is a snapshot only and not an observation of the evolution. In this situation it is appropriate to make use of the duality of the voter model to (instantaneously) coalescing random walks (which holds since our kernels $a(\cdot, \cdot)$ are symmetric). The state of the voter model at *one fixed* time of observation can be described in terms of a system of rate 1 coalescing random walks on S with jump distribution $a(\cdot, \cdot)$ (see e.g. Liggett (1985)).

The explicit procedure is the following. We start with a labelled particle at each site $x \in S$. In each step the following happens.

- An occupied site $x \in S$ is chosen at random.
- A "neighbour" y is chosen at random according to $a(\cdot, \cdot)$.
- All particles located at x are moved to y.
- All particles at y (if any) stay at y.

These steps are made repeatedly until the time of observation. (Actually the system time is not the number of step but rather the sum of $(\#\{x \in S : x \text{ is occupied}\})^{-1}$ along the steps.) Then (independently) we associate with each occupied site $x \in S$ a random colour $c(x) \in \{\text{White}, \text{Black}\}$. The *initial* positions of all particles located *now* at x get coloured $c(x)$.

Sample Paths of Diffusion

We make the naive approach to simulate the trajectories of the solution of the following stochastic differential equation

$$dX_t = a(X_t)\, dW_t\,, \qquad X_0 = x_0 \tag{A.1}$$

(with $a(x)$ the diffusion coefficient and (W_t) a standard Brownian motion).

We discretise the time and sum up small Brownian increments weighted with $a(X_t)$.

More precisely we fix $\Delta t > 0$ and sample an independent family $\widetilde{W}_0, \widetilde{W}_1, \widetilde{W}_2, \ldots$ of standard normally distributed random variables. We let $\widetilde{X}_0 := x_0$ and define successively

$$\widetilde{X}_{n+1} = \widetilde{X}_n + a(\widetilde{X}_n) \cdot \sqrt{\Delta t}\, \widetilde{W}_n.$$

Then $(\widetilde{X}_{[t/\Delta t]})_{t \geq 0}$ is an approximate trajectory.

Bibliography

[1] **Arratia, R. (1982)** Coalescing Brownian motion and the voter model on $\mathbb{Z}$. Unpublished manuscript

[2] **Baillon, J.-B., Clément, Ph., Greven, A., 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. Canadian Journal of Mathematics **47(1)**, 3-27

[3] **Bramson, M., Cox, J.T., Greven, A. (1993)** Ergodicity of Critical Spatial Branching Processes in Low Dimensions. Ann. Probab. **21**, 1946-1957

[4] **Bramson, M., Cox, J.T., Greven, A. (1995)** Invariant Measures in Critical Spatial Branching Processes in High Dimensions, Ann. Prob. (to appear)

[5] **Bramson, M., Griffeath, D. (1980)** Clustering and dispersion rates for some interacting particle systems on $\mathbb{Z}$. Ann. Probab. **8**, 183-213

[6] **Cassandro, M., Galves, A.Olivieri, E., Vares, M.E. (1984)** Metastable behavior of stochastic dynamics: a pathwise approach. J.Stat. Phys. **35**, 606-628

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

[8] **Cox, J.T., Dawson, D.A., Greven, A. (1995)** Universality properties for the historical process of state dependent branching and resampling systems. (Manuscript in preparation)

[9] **Cox, J.T., Durrett, R. (1994)** Hybrid zones and voter model interfaces. Manuscript in preparation.

[10] **Cox, J.T., Fleischmann, K., Greven, A. (1995)** Comparison of interacting diffusions and applications to their ergodic theory. Prob. Th. Rel. Fields (to appear).

[11] **Cox, J.T., Greven, A. (1990)** On the longterm behavior of some finite particle systems. Probab. Th. Rel. Fields **85**, 195-237

[12] **Cox, J.T., Greven, A. (1991)** On the longterm 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. Eds. R.Durrett and H.Kesten. Progress in Probab.**28**, 203-213, Birkhäuser, Boston.

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

[14] **Cox, J.T., Greven, A. (1994b)** The finite systems scheme: An abstract theorem and a new example. In D. Dawson (ed.) Measure-Valued Processes, Stochastic Partial Differential Equations and Interacting Systems, CRM Proc. Lecture Notes and Monographs **5**, pp. 55-67. Providence, RI: American Mathematical Society.

[15] **Cox, J.T., Greven, A., Shiga, T. (1994)** Finite and infinite systems of interacting diffusions. To appear in Probab. Th. Rel. Fields.

[16] **Cox, J.T., Griffeath, D. (1983)** Occupation time limit theorems for the voter model. Ann. Probab. **11(4)**, 876-893

[17] **Cox, J.T., Griffeath, D. (1985)** Occupation time limit theorems for critical branching brownian motions. Ann. Probab. **13(4)**, 1108-1132

[18] **Cox, J.T., Griffeath, D. (1986)** Diffusive clustering in the two dimensional voter model. Ann. Probab. **14(2)**, 347-370

[19] **Dawson, D. (1977)** The Critical Measure Diffusion. Z. Wahr. verw. Geb. **40**, 125-145

[20] **Dawson, D. (1993)** Measure-Valued Markov Processes. In: Ecole d'Eté de Probabilités de St.Flour XXI - 1991, LNM **1541**, Springer-Verlag.

[21] **Dawson, D.A., Gärtner, J. (1988)** Long time behavior of interacting diffusions. In: Stochastic Calculus in Application, Proceedings of the Cambridge Symposium, 1987, ed. J.R. Norris, Pitman Research Notes in Mathematics Volume 197, pages 29-54.

[22] **Dawson, D.A., Greven, A. (1993a)** Multiple time scale analysis of hierarchically interacting systems, A Festschrift to honor G. Kallianpur, 41-50. Springer, New York

[23] **Dawson, D.A., Greven, A. (1993b)** Multiple time scale analysis of interacting diffusions. Prob. Theory Rel. Fields **95**, 467-508

[24] **Dawson, D.A., Greven, A. (1995)** Hierarchically interacting Fleming-Viot processes with selection and mutation: Multiple time scale analysis and quasi equilibria. (Manuscript in preparation)

[25] **Dawson, D.A., Greven, A., Vaillancourt, J. (1995)** Equilibria and quasi equilibria for infinite systems of interacting Fleming-Viot processes. Trans. Am. Math. Soc. Vol. **347**, No. 7, pp 2277-2361

[26] **Donnelly, P., Welsh, D. (1983)** Finite particle systems and infection models. Math. Proc. Camb. Philos. Soc. **94**, 167-182

[27] **Durrett, R. (1979)** An Infinite Particle System with Additive Interactions, Adv. Appl. Prob. **11**, 355-383

[28] **Durrett, R., Liu, X. (1988)** The contact process on a finite set. Ann. Probab. **12**, 999-1040

[29] **Durrett, R., Schonmann, R.H. (1988)** The contact process on a finite set, II. Ann. Probab. **16**, 1570-1583

[30] **Durrett, R., Schonmann, R.H., Tanaka, N.I. (1989)** The contact process on a finite set: The critical case, III. Ann. Probab. **17**, 1303-1321

[31] **Erdös, P., Taylor, S.J. (1960)** Some problems concerning the structure of random walk paths. Acta Mathem. Acad. Sci. Hungar. **11**,137-162

[32] **Ethier, S.N., Kurtz, T.G. (1986)** Markov processes, characterization and convergence. Wiley, New York

[33] **Feller, W. (1966)** An Introduction to Probability Theory and Its Applications, Vol. 2. Wiley, New York, London, Sidney

[34] **Fleischman, J. (1978)** Limiting Distributions for Branching Random Fields. Trans. Am. Soc. Vol. **239**

[35] **Fleischmann, K., Greven, A. (1994)** Diffusive clustering in an infinite system of hierarchically interacting diffusions. Probab. Th. Rel. Fields. **98**, 517-566

[36] **Gorostiza, L.G., Roelly, S., and Wakolbinger, A. (1992)** Persistence of critical multitype particle systems and measure branching processes. Prob. Th. Rel. Fields **92**, 313-335

[37] **Gorostiza, L.G., Wakolbinger, A. (1991)** Persistence Criteria for a Class of Critical Branching Particle Systems in Continuous Time. Ann. Probab. **19**, 266-288

[38] **Gorostiza, L.G., Wakolbinger, A. (1993)** Long Time Behaviour of Critical Branching Particle Systems and Applications. In D. Dawson (ed.) Measure-Valued Processes, Stochastic Partial Differential Equations and Interacting Systems, CRM Proc. Lecture Notes and Monographs **5**, pp. 119-137. Providence, RI: American Mathematical Society.

[39] **Iscoe, I. (1986)** A weighted occupation time for a class of measure-valued branching processes, Prob. Th. Rel. Fields **71**, 85-116

[40] **Kallenberg, O. (1983)** Random measures, Akademie Verlag and Academic Press.

[41] **Kemeny, J.G., Snell, J.S., Knapp, A.W. (1976)** Denumerable Markov chains, 2. ed. Springer, New York

[42] **Klenke, A. (1995)** Different Clustering Regimes in Systems of Hierarchically Interacting Diffusions. Ann. Probab. (to appear)

[43] **Kopietz, A. (1995)** Clusterformationen bei wechselwirkenden Diffusionen. Diplomarbeit, Göttingen.

[44] **Lee, T.-Y. (1991)** Conditional Limit Distributions of Critical Branching Brownian Motions Ann. Probab. **19**, 289-311

[45] **Le Gall, J.-F. (1991)** Brownian Excursions, Tres and Measure-Valued Branching Processes Ann. Prob. **19**, 1399-1439

[46] **Liggett, T.M. (1985)** Interacting particle systems. Springer, NewYork

[47] **Rogers, L.C.G., Willams, D. (1987)** Diffusions, Markov processes and Martingales. Vol. 2: Itô calculus. John Wiley & Sons, New York.

[48] **Sawyer, S. (1976)** Results for the stepping stone model for migration in population genetics. Ann. Probab. **4**, 699-728

[49] **Schonmann, R.H. (1985)** Metastability for the contact process. J. Stat. Phys. **41**, 445-464

[50] **Shiga, T. (1980)** An interacting system in population genetics. J. Math. Kyoto Univ. **20-2**, 213-242

[51] **Shiga, T., Shimizu, A. (1980)** Infinite dimensional stochastic differential equations and their applications. J. Math. Kyoto Univ. **20-3**, 395-416

[52] **Spitzer, F. (1964)** Principles of random walk. Van Nostrand, Princeton, NJ

[53] **Tribe, R. (1993)** The long term behaviour of a stochastic PDE. Preprint No.74, Berlin.

[54] **Winter, A. (1995)** Clusterformationen bei wechselwirkenden Feller-Diffusionen. Diplomarbeit, Berlin.

[55] **Yamada, T., Watanabe, S. (1971)** On the uniqueness of solutions of stochastic differential equations. J. Math. Kyoto **11-1**, 155-167

List of Figures