Institut für Angewandte Analysis
und Stochastik

im Forschungsverbund Berlin e.V.

Diffusive clustering in an infinite system of hierarchically interacting diffusions

K. Fleischmann[1] , A. Greven[2]

submitted: 26th November 1992

[1] Institut für Angewandte Analysis
und Stochastik
Hausvogteiplatz 5-7
D – O 1086 Berlin
Germany

[2] Institut für Mathematische Stochastik
Lotzestraße 13
D – W 3400 Göttingen
Germany

Preprint No. 23
Berlin 1992

1991 Mathematics Subject Classification. Primary 60 K 35; secondary 60 J 60, 60 J 15.
Key words and phrases. Interacting diffusions, coalescing random walk, hierarchical system, cluster formation, universality.

Herausgegeben vom
Institut für Angewandte Analysis und Stochastik
Hausvogteiplatz 5-7
D – O 1086 Berlin

Fax: + 49 30 2004975
e-Mail (X.400): c=de;a=dbp;p=iaas-berlin;s=preprint
e-Mail (Internet): preprint@iaas-berlin.dbp.de

Diffusive Clustering in an Infinite System of Hierarchically Interacting Diffusions

By Klaus Fleischmann and Andreas Greven

Institut für Angewandte Analysis und Stochastik, Berlin

Institut für Mathematische Stochastik, Göttingen

Abstract. We study a system of linearly interacting diffusions on the interval $[0,1]$, indexed by an infinite hierarchical group with parameter N, modeling the evolution of gene frequencies accounting for migration and resampling. *Fisher-Wright diffusions* are the typical example for the diffusion term. A particular choice of the migration term guaranties that we are in the so-called diffusive clustering regime. The processes in the whole class considered and starting with a shift-ergodic initial law have qualitative properties just as in the Fisher-Wright case *(universality)*. Also, the initial law plays here no role. *Clusters* of components with values either close to *0* or close to *1* grow on various different scales *(diffusive clustering)*. More precisely, at time $N^t \to \infty$ (spatial) cluster sizes measured in a hierarchical distance grow like αt (i.e. consist of $N^{\alpha t}$ components) where $1-\alpha \in (0,1)$ is the hitting time of the traps $\{0,1\}$ for a time transformed Fisher-Wright diffusion. Nevertheless, the single components *oscillate* infinitely often between values close to *0* and close to *1*, but in such a way that they spend *fraction one* of their time together close to the boundary. Diffusive clustering phenomena we believe are in fact common to many interacting systems and were first discussed by Cox and Griffeath (1986) for the planar simple voter model. On the way, we prove some scaling results on a specific *coalescing random walk* with delay on the hierarchical group.

AMS 1980 subject classifications. Primary 60K35; secondary 60J60, 60J15

Key words and phrases. Interacting diffusions, coalescing random walk, hierarchical system, cluster formation, universality

1. Introduction and Main Results

1.a) Background and Motivation

A common feature of several systems with infinitely many interacting components and one conserved quantity (like the particle density) is the fact that one observes two types of longterm behavior depending on the interaction strength. In the case of a *"strong" interaction* we have a one parameter set of nontrivial extremal equilibrium states, one for each possible value of the conserved quantity. Furthermore, for reasonable initial states, the system approaches one of these equilibria as time tends to infinity. In the case of a *"weak" interaction* the only extremal invariant distributions are the *traps* of the system. This second situation is particularly interesting if we have at least two such traps. Then the system will *cluster*, i.e. for sets of components growing in time it will look locally like either one of the traps. Examples are the voter model, branching random walks, or systems of interacting diffusions. As a general reference to interacting particle systems, we refer to Liggett (1985).

In the case of a clustering system, one is of course interested in a more detailed description of the *clustering in time* as well as *in space:* (i) one would like to show that the components spend most of their time *"close to the traps"* but nevertheless *keep oscillating* infinitely often; (ii) one would like to know at which rate clusters grow in space as time tends to infinity. The latter depends on the strength of the interaction. Particularly interesting phenomena are observed in the *critical situation* where we have a weak interaction but close to the case of a strong interaction.

For example in Cox and Griffeath (1983, 1986), Bramson et al. (1986) the simple voter model has been studied, respectively in Cox and Griffeath (1990) the planar stepping stone model. (The simple voter model describes the random evolution of opinions, denoted by 0 or 1, of a community located on the sites of $\mathbb{Z}^d$, where each individual changes its opinion at a rate proportional to the number of neighbors who disagree. In the stepping stone model, more general opinions $0,1,\dots,r$ are allowed.) In the voter model, the complete consensus $\{\mathbf{x}\equiv 0$ or $\mathbf{x}\equiv 1\}$ constitutes the traps. For the critical dimension $d=2$ it was found that as time t tends to infinity clusters of all ones and all zeros grow like $t^{\alpha/2}$ where $\alpha\in(0,1)$ is a *random variable* related to a time transformed standard Fisher-Wright diffusion. This means that one observes *clusters* of various sizes which are as $t\to\infty$ of asymptotically *different orders of magnitude.* This is particularly interesting since it is visible already in simulations of very large but finite systems; see Cox and Greven (1991a).

Our first motivation is now to demonstrate that this phenomenon of *"diffusive clustering"* is not restricted to interacting systems of the type of the voter model but is universal, i.e. occurs in other systems of interest in applications, namely in a *whole class* of systems of interacting diffusions. Furthermore, this phenomenon does not depend on the initial distribution being a product measure. Even the Euclidean structure of the index set labelling the components is not relevant. For example, already in Dawson and Greven (1992a) such phenomenon was obtained at least in the *mean field limit* for a model of interacting diffusions on a hierarchical structure.

In the present paper we shall study some interacting diffusions on a hierarchical structure (see Definition 1.6 below) *without* passing to the mean field limit. This brings us to our second set of motivations: The model is interesting for reasons of application and from a theoretical point of view.

First, the particular case of interacting Fisher-Wright diffusions indexed by a hierarchical structure represents an interesting model in mathematical biology if one studies gene frequencies of individuals, with evolution accounting for *migration* and *resampling*, in the diffusion limit. Here an hierarchical notion of "neighboring" of components is more appropriate which takes into account only whether individuals belong to the same family, same clan, same village, same tribe etc., rather than taking into account some Euclidean notion of distance. See Sawyer and Felsenstein (1983), Dawson and Greven (1992b) for more references.

Secondly, in Dawson and Greven (1992a,b) it was shown in the mean field limit, that the growth of clusters in various scales is a phenomenon which is *"universal"* in a whole class of evolutions of the type of interacting diffusions. This universality is due to a *fixed point property* of Fisher-Wright systems under a renormalization procedure adapted to the *multiple time scale behavior* of such system. To make that rigorous it was important to have a hierarchical interaction and to take the mean field limit. Of course, this work raises the question what happens if one does not pass to the mean field limit. We will attack this problem by dealing first with interacting Fisher-Wright diffusions, where *duality* relations with coalescing random walks with delay are available, and then we shall generalize via *comparison arguments* (for such methods see also Cox and Greven (1991b)). One would hope to eventually recover universality here also from a fixed point property of the Fisher-Wright system, but this seems beyond the present technology.

In this paper we focus on the regime of diffusive clustering and we do not touch the quite interesting phenomenon arising in the formation of clusters in other regimes, i.e. fast and slow growing clusters. (For the mean field limit case, see again Dawson and Greven (1992a,b).)

Similar results can be expected for interacting diffusions indexed by $\mathbb{Z}^d$, but a more interesting direction of research would be to study different type models, for example, refine the known results on "random time averages" of spatial branching models in the critical dimension (see Cox and Griffeath (1985), Fleischmann and Gärtner (1986) and Iscoe (1986), or to study interacting Feller's branching diffusions in the critical dimension in order to find clusters of empty and crowded spots of different orders of magnitude. A larger sample of examples would hopefully reveal general principles ruling the intrinsic structure of the cluster formation in critical situations.

1.b) The Model

Start by introducing a countably infinite *hierarchical set* Ξ which shall label the components of our interacting system. Namely let Ξ denote the collection of all sequences $\xi=[\xi_1,\xi_2,\ldots]$ with coordinates ξ_i in the finite set $\{0,\ldots,N-1\}$ which are different from zero only finitely often. The natural number $N\geq 2$ is a once and for all *fixed parameter* of the model. Define the *"length"* (discrete norm) of ξ:

(1.1) $$\|\xi\| := \max\{i;\xi_i\neq 0\}, \quad \text{to be read as } 0 \text{ if } \xi=0=[0,0,\ldots].$$

One should think about this according to the following interpretation appearing in applications to mathematical biology. If a component has the label $\xi=[\xi_1,\xi_2,\ldots]$ then it is the ξ_1-st member of a family, which is the ξ_2-nd family of a clan, which is the ξ_3-rd clan of a village, which is the ξ_4-th village of a county, ..., which is the ξ_k*-th member of a* k*-level set*, $k\geq 1$. Hence, in the hierarchical lattice we think of components as grouped on different levels.

We regard Ξ as an *Abelian group* by defining the addition coordinate wise modulo N. Note that with $\|\xi-\zeta\|$, $\xi,\zeta\in\Xi$, we get a *hierarchical distance:* $\|\xi-\zeta\|=k$ means that the components with labels ξ and ζ are *"relatives of degree* k*"*. In other words, $\|\xi-\zeta\|$ is the coarsest level to separate the components with labels ξ and ζ.

Our system with interacting components will be constructed as random element in $(\mathbf{C}[\mathbb{R}_+,[0,1]])^{\Xi}$; i.e. with the component with label ξ we shall associate a continuous trajectory $x_\xi := \{x_\xi(t); t\geq 0\}$ in $\mathbf{C}[\mathbb{R}_+,[0,1]]$. Then setting $\mathfrak{X}(t) := (x_\xi(t))_{\xi\in\Xi}$ we obtain a *continuous trajectory* with values in the compact space $[0,1]^\Xi$ (which is always equipped with its product topology and with the Borel σ-field).

We start with describing the *evolution* of the desired stochastic process in an informal and heuristic fashion. The evolution of the system $\mathfrak{X}:=(x_\xi)_{\xi\in\Xi}$ will be based on two effects: on migration and various types of resampling. Let $\mathcal{G}^0$ denote the set of all functions g satisfying

(1.2) $\quad g\colon [0,1]\mapsto\mathbb{R}_+$ is Lipschitz continuous with $g(0)=0=g(1)$ and $g>0$ otherwise.

To model the *resampling effect*, we let each component x_ξ perform independently a *diffusion in* $[0,1]$ *with diffusion coefficient* $g\in\mathcal{G}^0$:

$$dZ(t) = \sqrt{g(Z(t))}\, dw(t), \qquad t>0,\ Z(0)\in[0,1],$$

where w is a one-dimensional standard Brownian motion. Note that such a diffusion has the *traps 0* and *1*, where we have to distinguish between two cases. If $\int_0^{1/2} dr\ r/g(r)$ and $\int_{1/2}^1 dr\ (1-r)/g(r)$ are both finite then the diffusion will end up in finite time in $\{0,1\}$. Otherwise the paths converge as $t\to\infty$ to the traps without ever hitting them. As a prototype for the first case we have the *Fisher-Wright diffusion:*

(1.3) $\quad g=b\ell$ where $\ell(r):=r(1-r)$, $0\leq r\leq 1$, and $b>0$ a *diffusion constant.*

An example for the second case which is interesting for applications as well is the *Ohta-Kimura diffusion:*

(1.4) $$g=b\ell^2 \quad\text{with}\quad b>0;$$

Ohta and Kimura (1973).

To model the *migration* effect, we add the following *hierarchical interaction*: The ξ-th component at time t (which has state $x_\xi(t)$) will get a *drift* defined by the deviation from the empirical means

(1.5) $$x_{\xi,k}(t) := N^{-k}\sum_{\zeta\in\Xi} 1\{\|\xi-\zeta\|\leq k\}\ x_\zeta(t), \qquad \xi\in\Xi,\ k\geq 1,\ t\geq 0,$$

of the states of all its relatives of at most degree k (*block average of level* k *around* ξ). The different block levels k will additionally be weighted by the factor αN^{-k} where α is a fixed positive constant. In this sense we have a *hierarchical local interaction*. This interaction forces the components to be in line with their neighbors. That particular choice αN^{-k} of the weights (exponential decay) ensures that we are in the *critical region*. This makes the model the analog of the *planar* simple voter model. (For more rapidly decreasing weights one has different patterns of cluster formations resembling more the one-dimensional simple voter model; compare Dawson and Greven (1992a).)

In the Fisher-Wright case, this model has for example the following *interpretation in mathematical biology:* $\mathfrak{x}_\xi$ describes the evolution of a *gene frequency* at a colony ξ in the diffusion limit of a large number of individuals per colony and slow migration rate: a new born individual has a gene chosen with a probability given by the empirical frequency in the previous generation in the colony and each individual can migrate to other colonies. (See Ethier and Kurtz (1986), Chapt. 10.)

Let $\mathfrak{P}$ denote the set of all probability laws on $[0,1]^{\Xi}$. We are ready now for a rigorous definition of our process.

Definition 1.6. The *system* $\mathfrak{X} := \{\mathfrak{X}(t);t\geq 0\} := (\{\mathfrak{x}_\xi(t);t\geq 0\})_{\xi\in\Xi} =: [\mathfrak{X},\mathbb{P}^{g}_{\mu},\mu\in\mathfrak{P}]$ *of interacting diffusions* in the interval $[0,1]$ with *diffusion coefficient* $g\in\mathcal{G}^0$ (defined in (1.2)) is the *unique strong solution* of the following system of stochastic differential equations

$$
\begin{aligned}
(1.7)\qquad d\mathfrak{x}_\xi(t) &= \alpha\Big(\sum_{k=1}^{\infty} N^{-k}(\mathfrak{x}_{\xi,k}(t)-\mathfrak{x}_\xi(t))\Big)dt + \sqrt{g(\mathfrak{x}_\xi(t))}\; dw_\xi(t)\\
\mathfrak{x}_{\xi,k}(t) &:= N^{-k}\sum_{\zeta\in\Xi} 1\{\|\xi-\zeta\|\leq k\}\; \mathfrak{x}_\zeta(t), \qquad \xi\in\Xi,\ k\geq 1,\ t\geq 0,\\
\mathcal{L}(\mathfrak{X}(0)) &= \mu\in\mathfrak{P},
\end{aligned}
$$

where $\mathbf{w}:=\{w_\xi(t);t\geq 0\}_{\xi\in\Xi}$ is a system of independent one-dimensional standard Brownian motions and the parameters $N\geq 2$, $\alpha>0$, and $g\in\mathcal{G}^0$ are fixed. ∎

The solution $\mathfrak{X}$ is a *time-homogeneous strong Markov process* with continuous paths.

To apply directly known results on the existence of a unique strong solution as well as for the interpretation of such system of interacting diffusions, we remark that (1.7) can alternatively be written as

$$(1.8) \qquad d\mathfrak{x}_\xi(t) = \sum_{\zeta:\zeta\neq\xi} q_{\xi,\zeta}(\mathfrak{x}_\zeta(t)-\mathfrak{x}_\xi(t))dt + \sqrt{g(\mathfrak{x}_\xi(t))}\, dw_\xi(t), \qquad \xi\in\Xi,\ t\geq 0,$$

where

$$(1.9) \qquad q_{\xi,\zeta} := a \sum_{k=1}^{\infty} N^{-k}\, N^{-k}\, 1\{1\leq\|\xi-\zeta\|\leq k\} = aN^{-2\|\xi-\zeta\|+2}/(N^2-1), \qquad \xi\neq\zeta.$$

Hence, the diffusions of different components ξ and ζ interact in a way, corresponding to a *migration from* ζ *to* ξ *at rate* $q_{\xi,\zeta}$. The *total migration rate* into ξ is given by

$$(1.10) \qquad -q_{\xi,\xi} := \sum_{\zeta:\zeta\neq\xi} q_{\xi,\zeta} \equiv aN/(N^2-1), \qquad \xi\in\Xi,$$

since ξ has $N^{k-1}(N-1)$ relatives ζ of degree $k\geq 1$.

Note that this system of stochastic differential equations has a unique strong solution, due to Theorem 3.2 in Shiga and Shimizu (1980).

As a rule, we shall study the case where the *initial state* $\mathfrak{X}(0)$ has a law in the set $\mathfrak{T}_\theta$ of all *shift-invariant* and *ergodic* distributions $\mu\in\mathfrak{P}$ with expectation $\mathbb{E}_\mu\mathfrak{x}_0(0) =: \theta$ where the overall *density* $\theta\in(0,1)$ is fixed once and for all. In particular, μ could be the product measure with marginals $\theta\delta_1 + (1-\theta)\delta_0$, or even the degenerate law $\mu=\delta_{\bar\theta}$ with constant state $\bar\theta$: $\bar\theta_\xi\equiv\theta$.

1.c) Review of the Basic Ergodic Theory

Let $g\in\mathcal{G}^0$, $\mu\in\mathfrak{T}_\theta$. Then the solution to (1.7) has the following *longterm behavior:*

$$(1.11) \qquad \mathbb{P}^g_\mu(\mathfrak{X}(t)\in(\cdot)) \underset{t\to\infty}{\Longrightarrow} (1-\theta)\delta_{\mathbf{0}} + \theta\delta_{\mathbf{1}}$$

where $\mathbf{0}$ and $\mathbf{1}$ are the constant states $\mathfrak{x}_\xi\equiv 0$ and $\mathfrak{x}_\xi\equiv 1$ (*"extreme consensus"*) in $[0,1]^\Xi$. (Here and in the sequel $\Longrightarrow$ denotes weak convergence of probability laws.) This follows from Cox and Greven (1991b), combined with the fact that related to a result of Sawyer and Felsenstein (1983) the continuous time *random walk* in Ξ with homogeneous and symmetric rates $q_{\xi,\zeta}$ for a jump from ξ to $\zeta\neq\xi$ (according to (1.9) and (1.10)) is *recurrent*; see Lemma 2.21 below. (By the way, in the transient case, instead of the "degeneration" (1.11) one has convergence to a non-trivial equilibrium state.)

Even though (1.11) implies

$$\mathbb{P}^g_\mu(\mathfrak{x}_\xi(t) \in [\varepsilon,1-\varepsilon]) \underset{t\to\infty}{\longrightarrow} 0, \qquad \xi\in\Xi,\ 0<\varepsilon<1/2,$$

note first of all that

$$\mathbb{P}_{\mu}^{q}(\mathfrak{x}_{\xi}(t) \in (0,1)) \equiv 1, \qquad \xi\in\Xi,\ t>0, \tag{1.12}$$

and second that (1.11) does say nothing about whether components oscillate. (To prove (1.12), just observe that, for each $a\in(0,1)$ and $q\in\mathcal{G}^{0}$, a diffusion of the form $d\mathfrak{x}_{\xi}(t) = (a-\mathfrak{x}_{\xi}(t))dt + \sqrt{q(\mathfrak{x}_{\xi}(t))}\, dw_{\xi}(t)$ has this property.) Nevertheless, for large time locally one has either a cluster of only values close to one or a cluster of only values close to zero.

The *main purpose of this paper* is to investigate how fast clusters will grow during the evolution. But we shall also answer the question where finite collections of components spend most of their time, and how their normalized occupation time behaves as $t\to\infty$. Before we will state the main results we give some of the history and the background of this problem.

In Dawson and Greven (1992a) it was shown that in the *mean field limit* $N\to\infty$ and at time scale N^{t} various regimes of clustering are possible: large clusters, diffusive clustering, and small clusters, corresponding to whether (for large N) cluster sizes grow, respectively, like t, like αt with a random variable $\alpha\in(0,1)$, or like a function f(t) which also tends to ∞ as $t\to\infty$, but as o(t). Here cluster sizes are measured by the hierarchical distance.

For analogy, as $t\to\infty$ the *simple voter model* in dimension one has large clusters namely ones of order $t^{1/2}$, in dimension two it clusters diffusive with diameters of order $t^{\alpha/2}$, where $\alpha\in(0,1)$ is random, whereas small clusters will not occur in the simple voter model since the three dimensional case is already "stable"; we refer to Cox and Griffeath (1986).

Another model exhibiting similar phenomena are *Dawson-Watanabe processes (superprocesses)* in $\mathbb{R}^{d}$: In subcritical dimensions d one has large clusters, in the critical dimension the diffusive clumping picture is reflected by a self-similarity property and by the random time average behavior, and in supercritical dimensions there exist already non-trivial steady states; see Dawson and Fleischmann (1988) and references therein. Note also Remark 1.26 below.

We shall focus here on the most interesting regime of *diffusive clustering*, but now (as already said) for *fixed parameter* N. Following the scheme in Cox and Griffeath (1986), we shall describe the growth of clusters from three different points of view, each one culminating in a theorem. Here we stress the interesting fact, that more or less we get the *same limit expressions* as there, although at the first sight the interacting diffusion model here with general diffusion coefficient looks rather different than the voter model. A fourth theorem is then devoted to the time picture of components.

1.d) Main Results

Our main results will be presented in four parts: three parts describing various aspects of the clustering in space, and one dealing with clustering viewed in time.

(i) Spatially Thinned Systems. In order to assess how far correlations between the components reach as $t\to\infty$, we will consider first some "thinned out systems". To define them, we introduce the *shift operators* $\{\mathcal{S}_m: m\geq 0\}$ in Ξ by setting $(\mathcal{S}_m\zeta)_i:=\zeta_{m+i}$, $\zeta\in\Xi$, $i\geq 1$. That is, $\mathcal{S}_m$ cuts off the first m coordinates (levels). Conversely, for each $m\geq 0$, we will understand the "inverse operator" $\mathcal{S}_m^{-1}$: $\Xi\mapsto\Xi$ as being *any* fixed function such that $\mathcal{S}_m\mathcal{S}_m^{-1}\xi=\xi$, $\xi\in\Xi$. (For instance, take always 0 for the coordinates to be newly created.) For fixed $\alpha,t\geq 0$ we now define a *thinned out (rescaled) system* ${}^{\alpha}\mathfrak{X}(t) := ({}^{\alpha}\mathfrak{x}_{\xi}(t))_{\xi\in\Xi}$ by setting

$$(1.13)\qquad {}^{\alpha}\mathfrak{x}_{\xi}(t) := \mathfrak{x}_{\zeta}(N^t) \quad \text{where} \quad \zeta = \mathcal{S}_{[\alpha t]}^{-1}\xi, \quad \xi\in\Xi.$$

Note that if ξ and ξ' have the distance $\|\xi-\xi'\|=k\geq 0$ then the ζ,ζ' associated by $\mathcal{S}_{[\alpha t]}^{-1}$ have the distance $\|\zeta-\zeta'\|=[\alpha t]+k=\alpha t+O(1)$ as $t\to\infty$. Roughly speaking, at time N^t we look only on such a thinned out system in which fixed collections of different labels have a distance of order $\alpha t + O(1)$. It will turn out that in the case $0<\alpha<1$ opposed to (1.11) extreme consensus $\{0,1\}$ will appear (as $t\to\infty$) only with a probability strictly between 0 and 1.

In order to give a precise description of this effect in the following theorem, recall first that for $\mu\in\mathfrak{I}_\theta$ the constant $\theta\in(0,1)$ is the initial density of $\mathfrak{X}$, and second introduce the following objects.

Definition 1.14. Let $Y^\theta := \{Y^\theta(s); 0\le s\le +\infty\}$ denote the *standard Fisher-Wright diffusion* starting at θ, that is the process solving

$$dY(t) = \sqrt{Y(t)(1-Y(t))}\, dw(t), \quad t\ge 0,\ Y(0)=\theta. \tag{1.15}$$

Recall that $Y^\theta(+\infty) \in \{0,1\}$. Set $\tilde{Y}(\alpha) := Y^\theta(-\log(\alpha\wedge 1))$, $\alpha\ge 0$, and define $F_{\theta,\alpha}$ as the law of the random variable $\tilde{Y}(\alpha)$. Moreover, for $\tilde{\theta}\in[0,1]$ let $\pi_{\tilde{\theta}}$ denote the *product measure* on $\{0,1\}^\Xi$ with marginals $(1-\tilde{\theta})\delta_0 + \tilde{\theta}\delta_1$. ▮

Theorem 1 (spatially thinned systems). *Let $\mu\in\mathfrak{I}_\theta$ and $q\in\mathcal{G}^0$. For each $\alpha\ge 0$, the law of the thinned out systems ${}^\alpha\mathfrak{X}(t)$ with respect to $\mathbb{P}^q_\mu$ converges to a mixture of product measures:*

$$\mathcal{L}({}^\alpha\mathfrak{X}(t)) \underset{t\to\infty}{\Longrightarrow} \int_0^1 F_{\theta,\alpha}(d\tilde{\theta})\, \pi_{\tilde{\theta}}. \qquad \bullet$$

Consequently, in all α cases, the limiting components take on only the values 0 or 1. In the "strongly" thinned out case $\alpha\ge 1$ the law $F_{\theta,\alpha}$ of $\tilde{Y}(\alpha)$ degenerates to δ_θ, hence the limiting field ${}^\alpha\mathfrak{X}(\infty)$, say, has independent components, i.e. all components belong to "independent clusters". On the other hand, for $\alpha=0$, that is $-\log\alpha=+\infty$, the "random" density $\tilde{\theta}$ has distribution $\theta\delta_1+(1-\theta)\delta_0$, that is the product law degenerates to δ_0 or δ_1, and (1.11) is recovered. If $0<\alpha<1$, extreme consensus $\{0,1\}$ will be met with a certain probability (which decreases in α), but also one observes fluctuating components, belonging to different clusters. Summarizing, at the exponential time scale N^t *correlations* between components are build up even over distances $[\alpha t]$ for all $\alpha\in(0,1)$, and we can control the strength of this correlation via the distribution $F_{\theta,\alpha}$ of the standard Fisher-Wright diffusion Y^θ taken at time $-\log(\alpha\wedge 1)$.

(ii) Time-dependent Block Averages. The next question is how averages over large blocks do behave as a function of the block size growing in time (recall (1.5)). We denote by $\overset{fdd}{\Longrightarrow}$ weak convergence of all finite dimensional distributions. Recall $\{\tilde{Y}(\alpha); \alpha\ge 0\}$, the time reversed Fisher-Wright diffusion on a logarithmic scale, see Definition 1.14.

Theorem 2 (time-dependent block averages). *Let* $\mu\in\mathcal{I}_\theta$, $\mathcal{q}\in\mathcal{G}^0$, *and* $\xi\in\Xi$. *With respect to* $\mathbb{P}_\mu^{\mathcal{q}}$, *the behavior of block averages is given by*

$$\mathcal{L}\Big(\{\mathcal{I}_{\xi,[\alpha t]}(N^t);\alpha\geq 0\}\Big) \underset{t\to\infty}{\overset{fdd}{\Longrightarrow}} \mathcal{L}\Big(\{\tilde{Y}(\alpha);\alpha\geq 0\}\Big). \tag{1.16}$$ •

This means, roughly speaking, that at time N^t (as $t\to\infty$), the block averages $\mathcal{I}_{\xi,[\alpha t]}(N^t)$ of "superlarge" levels $[\alpha t]$, $\alpha\geq 1$, approach the deterministic value $Y^\theta(0)=\theta$ which is the initial density of $\mathfrak{X}$; that is, empirical means are built over various clusters resulting into the system mean θ. But at each level of order αt, $0<\alpha<1$, the block averages $\mathcal{I}_{\xi,[\alpha t]}(N^t)$ (over blocks of $N^{[\alpha t]}$ components) *remain random* also in the limit and will *fluctuate* with α according to $\tilde{Y}(\alpha) = Y^\theta(\log(1/\alpha))$. For averages over blocks of "small" level $o(t)$, in the limit we see either the value 0 or 1. Note that for fixed $\alpha\in(0,1)$ and in the limit $t\to\infty$, with a certain positive probability these block averages will be 0 or 1 (where 0 and 1 are the traps of Y^θ), and the probability of that extreme consensus will increase from 0 to 1 as α varies from 1 to 0. In other words, the empirical means stabilize in distribution at separated scales in different ways and, in particular, exhibit the *diffusive clustering phenomenon:* clusters of values close to 0 and close to 1 grow on a *variety of spatial scales.*

Remark 1.17. Opposed to the mean field limit case [11], but as in [8], it remains an *open problem* whether or not convergence in path space of continuous functions occurs in Theorem 2 (i.e. *tightness* holds). □

(iii) Cluster Sizes. In order to describe the phenomenon of diffusive clustering in another and more "geometric" way, we introduce the following notations. For $t\geq 0$ and $0<\varepsilon<1/2$, denote by $\mathcal{B}_t^\varepsilon$ the set of all configurations $z\in[0,1]^\Xi$ such that the *sum* over all components z_ξ with $\|\xi\|\leq t$ is either smaller than ε or bigger than $N^t-\varepsilon$. Roughly speaking, $\mathcal{B}_t^\varepsilon$ describes the event that the block $\{\xi\in\Xi;\|\xi\|\leq t\}$ around 0 of relatives of at most degree $[t]$ is covered by an "*ε-cluster*" that is, in particular, no component is more than ε away from 0 or 1, and even the total deviation is less than ε. Define the random variables

(1.18) $$S_t^{\varepsilon} := \sup\{\alpha>0;\ \mathfrak{X}(N^t)\in\mathfrak{B}^{\varepsilon}_{\alpha t}\} \quad \text{and} \quad T := \sup\{\alpha>0;\ \tilde{Y}(\alpha)\in\{0,1\}\}$$

in $[0,+\infty]$ (to read as 0 if the corresponding set is empty). S_t^{ε} describes the *normalized size* of the largest block covered by an ε-cluster, whereas $1-T$ is the *hitting time of the traps* of a time transformed standard Fisher-Wright diffusion (note that $0<T<1$ a.s.). The following theorem will only be proved up to a statement on coalescing random walks with delay (see *Hypothesis* 7.2 in Section 7), proven in the analogous case of simple coalescing random walks on $\mathbb{Z}^2$ in Bramson et al. (1986).

Theorem 3 (cluster sizes). *Let* $\mu\in\mathfrak{T}_{\theta}$*. In the Fisher-Wright case* $g=b\ell$*,* $b>0$*, (as described in* (1.3)*), the normalized* ε*-cluster sizes* S_t^{ε} *converge in law with respect to* $\mathbb{P}_{\mu}^{b\ell}$ *towards* T*:*

$$\lim_{\varepsilon\to 0}\lim_{t\to\infty} \mathcal{L}(S_t^{\varepsilon}) = \mathcal{L}(T). \qquad \bullet$$

Loosely speaking, even in the limit as $t\to\infty$ the blocks completely covered by an ε-cluster at time N^t have a *random* size of order Tt, i.e. T is the *random velocity of growth for* ε*-clump sizes*. Note that the *asymptotic "volume"* of the block clusters at time N^t is given by N^{tT}.

(iv) Time Picture of Components. Finally we will ask how the components $\mathfrak{x}_{\xi}$ behave *through time.*

Theorem 4.a) (clumps sticking at the boundary). *Consider any* $\mu\in\mathfrak{P}$ *and* $g\in\mathcal{G}^0$*. For each finite collection* $\xi_1,\dots,\xi_m\in\Xi$ *and* $0<\delta<1/2$*,*

(1.19) $$t^{-1}\int_0^t ds\ 1\left\{\mathfrak{x}_{\xi_1}(s),\dots,\mathfrak{x}_{\xi_m}(s) < \delta \ \ or \ \ \mathfrak{x}_{\xi_1}(s),\dots,\mathfrak{x}_{\xi_m}(s) > 1-\delta\right\} \xrightarrow[t\to\infty]{} 1$$

in $\mathbb{P}_{\mu}^{g}$*-probability. Relation* (1.19) *is true even* $\mathbb{P}_{\mu}^{g}$*-a.s. provided that the following additional Condition* 1.20 *on* g *or* μ *holds.* $\bullet$

Condition 1.20. *Suppose*

(1.21) $$\liminf_{r\to 0} g(r)/r > 0 \quad \text{and} \quad \liminf_{r\to 1} g(r)/(1-r) > 0$$

(roughly speaking, g does not have a vanishing derivative at the boundary), *or*

(1.22) $$\mu([\varepsilon,1-\varepsilon]^{\Xi}) = 1 \quad \text{for some} \quad 0 < \varepsilon < 1/2$$

(that is, initially all components are bounded away from *0* and *1*). ∎

Note that (1.21) is valid in the Fisher-Wright case (1.3) but fails to hold for the Ohta-Kimura diffusion (1.4).

Remark 1.23. We have no doubts that the Condition 1.20 is only of a technical nature and could be dropped. However, the methods used so far in the present paper seem to be insufficient to get this out smoothly. □

Theorem 4.a) says that each fixed finite collection of components *sticks together "most of the time" close to a boundary point* of the interval $[0,1]$. On the other hand, the components actually *oscillate a.s.,* since a *law of large numbers* holds:

Theorem 4.b) (law of large numbers and oscillations). *For* $\mu\in\mathcal{I}_\theta$, $g\in\mathcal{G}^0$, $\xi\in\Xi$,

(1.24) $$t^{-1}\int_0^t ds\ \mathfrak{x}_\xi(s) \underset{t\to\infty}{\longrightarrow} \theta \quad \textit{in } \mathbb{P}_\mu^g\textit{-probability,}$$

(1.25) $$\limsup_{t\to\infty} \mathfrak{x}_\xi(t) = 1 \quad \textit{and} \quad \liminf_{t\to\infty} \mathfrak{x}_\xi(t) = 0, \qquad \mathbb{P}_\mu^g\text{-a.s.} \qquad \bullet$$

Remark 1.26. Spatially homogeneous branching processes in the critical dimension behave similar in that "components" keep oscillating between *0* and $+\infty$. But they are different in that (formally) the empty population is the single trap (and there is no other steady state, see Bramson et al. (1992)). Moreover, as already mentioned above, opposed to the LLN (1.24), averages remain random. □

1.e) Outline

A principal tool for the proofs of the theorems above is the fact that in the special case of *interacting Fisher-Wright diffusions* (1.3), i.e. if $g=b\mathfrak{f}$, a *dual process* η exists (Shiga (1980)). This is a system of *coalescing random walks with delay* on the countably infinite Abelian group Ξ, where the random walk is determined by the migration term in (1.7) whereas the interaction (coalescence) is related to the diffusion term. The duality relation simplifies even further if $\mathfrak{X}(0)$ is distributed according to a product measure. We then use *moment calculations, coupling and comparison arguments* to deduce the theorems for the general case, i.e. to get *universality* in μ and g.

In Section 2 we shall start studying the asymptotics of *hitting times* for random walks on the hierarchical group. In Section 3 we proceed to systems η of coalescing random walks with delay and approximate them by systems

$\tilde{\eta}$ without delay, i.e. by ordinary coalescing random walks. For the latter we prove a scaling limit proposition (as $t\to\infty$) concerning the probabilities finding exactly k particles at time N^t if we start initially with n particles, but which are at an hierarchical distance growing according to a linear function of t (see Proposition 3.13). The Theorems 4.a) and 4.b) will be verified in Section 4. The first one will then be used for the proofs of the Theorems 1-3 in the remaining Sections 5-7, respectively.

2. Preparations: On Random Walks on the Hierarchical Group

This section presents systematically all the facts on random walks on the hierarchical group which are needed later in the proofs. We begin with collecting in 2.a) and b) some facts on random walks on the hierarchical group Ξ. For the Lemmas 2.2, 2.7, 2.9, and (the discrete time analog of) Lemma 2.21, compare Sawyer and Felsenstein (1983). In 2.c) and e) we present the asymptotics for the law of the *first hitting time* of the "origin" 0 of Ξ (see the Propositions 2.37 and 2.43 below). The particular random walk we are studying is an analog of the planar simple random walk in the sense that the distribution of the return time to 0 has tails of order $1/\log t$ as $t\to\infty$. In Subsection 2.f) and g) we consider *collision times* of systems of *independent* such walks. As an application, an asymptotics of some moments of the interacting diffusion system are derived in the final subsection.

2.a) A Discrete Time Random Walk: Iterates of the Transition Kernel

We start with some prerequisites. First we shall introduce the notion of a characteristic function of a law on Ξ and recall its properties. The argument of such a characteristic function will run in the "dual" set $\Xi^* := \{[y_1,y_2,\ldots];\ y_i\in\{0,1,\ldots N-1\},\ i\geq 1\}$ equipped with the product topology. This set can be decomposed as follows: $\Xi^* = \{0\}\cup\bigcup_{m=1}^{\infty}\Xi^*_m$. Here Ξ^*_m consists of those elements y of Ξ^* with the property that $y_1=\ldots=y_{m-1}=0$ but $y_m\neq 0$. Defining the addition + again coordinate wise modulo N, this Ξ^* is an uncountable Abelian group, the *group of characters.*

The *uniform distribution* (normed Haar measure) h on (the Borel σ-field of) the compact space Ξ^* can be identified via the following three conditions (to see this, note that Ξ^*_m has "N times more elements" than Ξ^*_{m+1}):

(i) Set $h(\{0\}) := 0$ and $h(\Xi^*_m) := (N-1)N^{-m}$ for $m\geq 1$.

(ii) The coordinates $y_m, y_{m+1}, \ldots$ of y are *independent* with respect to the conditional distribution $\hbar(\cdot\,|\,\Xi^*_m)$, $m \geq 1$.

(iii) With respect to this conditional distribution, y_m is *"uniformly"* distributed on the finite set $\{1,\ldots,N-1\}$ whereas all other $y_{m+1}, y_{m+2}, \ldots$ are *"uniform"* on $\{0,1,\ldots,N-1\}$.

Now we are ready to define the *characteristic function* $\hat{p}$ of a probability law p on (the Borel σ-field of) Ξ. We shall use the notation $(\xi,y) := \sum_{j\geq 1} \xi_j y_j$.

Set:

$$\hat{p}(y) := \sum_{\xi\in\Xi} p_\xi \, e^{2\pi i(\xi,y)/N}, \qquad y\in\Xi^*.$$

Note that $|\hat{p}(y)| \leq 1 = \hat{p}(\mathbf{0})$.

In the sequel we shall need the explicit form of the Fourier transform of a particular distribution on Ξ, as well as a representation of the n-th convolution power of this distribution by inverse Fourier transforms. As an intermediate step we derive the *inversion formula*

$$p_\xi = \int \hbar(dy)\, \hat{p}(y)\, e^{-2\pi i(\xi,y)/N}, \qquad \xi\in\Xi. \tag{2.1}$$

The proof is by explicit calculation using the following formula.

Lemma 2.2. *For each* $\xi\in\Xi$: $\displaystyle\int \hbar(dy)\, e^{2\pi i(\xi,y)/N} = \delta_0(\xi).$

Proof. The set Ξ can be represented as $\Xi = \bigcup_{k\geq 0} \Xi_k$ with $\Xi_k := \{\xi\in\Xi; \|\xi\|=k\}$. The case $\xi=\mathbf{0}$ is trivial. For $\xi\in\Xi_k$, $k>0$, we have $\xi_i=0$ for all $i>k$, and the integral in the lemma can be written as

$$= \sum_{m\geq 1} (N-1)N^{-m} \int \hbar(dy\,|\,\Xi^*_m) \exp\Big[2\pi i \sum_{j:m\leq j\leq k} \xi_j y_j/N\Big].$$

Because of $\xi_k \neq 0$, the sum formula for a finite geometric series yields

$$\sum_{y_k=0}^{N-1} \exp[2\pi i\xi_k y_k/N] = (1 - \exp[2\pi i\xi_k])/(1 - \exp[2\pi i\xi_k/N]) = 0, \tag{2.3}$$

since ξ_k is in particular an integer. We get

$$\int \hbar(dy\,|\,\Xi^*_m) \exp\Big[2\pi i \sum_{j:m\leq j\leq k} \xi_j y_j/N\Big] = \begin{cases} 0 & \text{if } m<k, \\ -1/(N-1) & \text{if } m=k, \\ 1 & \text{if } m>k, \end{cases} \qquad \xi\in\Xi_k. \tag{2.4}$$

Hence the claim follows. ■

The distributions we will be interested in have the property that p_ξ only depends on $\|\xi\|$. More precisely, let r be any distribution on $\mathbb{Z}_+ :=$

$\{0,1,\ldots\}$. Then define an associated probability law p on Ξ by setting

(2.5) $$p_\xi := r_{\|\xi\|}/R_{\|\xi\|}, \qquad \xi\in\Xi,$$

where

(2.6) $$R_k := N^{k-1}(N-1) + \delta_0(k)/N, \qquad k\geq 0,$$

is the number of elements in the *k-th level set* $\Xi_k=\{\xi\in\Xi; \|\xi\|=k\}$. Since such p is *"uniform"* on each Ξ_k its characteristic function is relatively simple:

Lemma 2.7. *For distributions p of the form* (2.5),

(2.8) $$\hat{p}(y) = r_0+\ldots+r_{m-1} - r_m/(N-1) =: f_m \quad \textit{for} \quad y\in\Xi^*_m, \quad m\geq 1.$$

Proof. Write

$$\hat{p}(y) = \textstyle\sum_{k\geq 0} r_k R_k^{-1} \sum_{\xi\in\Xi_k} \exp\left[2\pi i \sum_{j:m\leq j\leq k} \xi_j y_j/N\right], \qquad y\in\Xi^*_m,\ m\geq 1,$$

use (2.3) and (2.6) to get

$$R_k^{-1} \sum_{\xi\in\Xi_k} \exp\left[2\pi i \sum_{j:m\leq j\leq k} \xi_j y_j/N\right] = \begin{cases} 1 & \text{if } k<m, \\ -1/(N-1) & \text{if } k=m, \\ 0 & \text{if } k>m, \end{cases} \qquad y\in\Xi^*_m. \qquad \blacksquare$$

Combining the inversion formula (2.1), Lemma 2.7, and (2.4) we conclude

Lemma 2.9. *For distributions p of the form* (2.5),

$$p_\xi = -[1-\delta_0(\xi)]\, N^{-\|\xi\|} f_{\|\xi\|} + (N-1)\textstyle\sum_{m:m>\|\xi\|} N^{-m} f_m, \qquad \xi\in\Xi,$$

with f_m *defined in* (2.8).

Now we use such law p of the form (2.5) as step distribution for a *discrete time random walk* on Ξ starting in $\mathbf{0}$. Then the law $p^{(n)}_{0,(\cdot)}$ of its state at time $n\geq 0$ has characteristic function $(\hat{p})^n$ and is again of the form (2.5). Thus we can apply the inversion formula in Lemma 2.9 to obtain

(2.10) $$p^{(n)}_{0,\xi} = -[1-\delta_0(\xi)]\, N^{-\|\xi\|} f^n_{\|\xi\|} + (N-1)\textstyle\sum_{m>\|\xi\|} N^{-m} f^n_m, \qquad n\geq 0,\ \xi\in\Xi.$$

2.b) Continuous Time Random Walk: Asymptotic Properties of the Transition Kernel

The purpose of this subsection is first to pass to a *continuous time* random walk $Z=\{Z(t);t\geq 0\}$ in Ξ with some *jump rate* $\kappa>0$ and "uniform" step probabilities as given in (2.5), and second to derive in a particular case several asymptotic results for the corresponding transition kernel as $t\to\infty$.

Let $p(t,\zeta,\xi)$, $t>0$, $\xi,\zeta\in\Xi$, denote the *transition probabilities* of this random walk Z, and P_ζ the law on $\mathbf{D}[\mathbb{R}_+,\Xi]$ of that process if it starts at site $\zeta\in\Xi$. From

$$p(t,0,\xi) = \sum_{n=0}^{\infty} e^{-\kappa t}\,\frac{(\kappa t)^n}{n!}\,p_{0,\xi}^{(n)}, \qquad t\geq 0,\ \xi\in\Xi,$$

and (2.10) we get

$$(2.11)\qquad p(t,0,\xi) = -\,[1-\delta_0(\xi)]\,N^{-\|\xi\|}\exp[-t\kappa(1-f_{\|\xi\|})]$$
$$+\,(N-1)\sum_{m>\|\xi\|} N^{-m}\exp[-t\kappa(1-f_m)],\qquad t\geq 0,\ \xi\in\Xi.$$

Hence,

$$(2.12)\qquad \sup_{\xi\in\Xi} p(t,0,\xi) = p(t,0,0),\qquad t\geq 0.$$

Recall the constant α from (1.7). *From now on we set*

$$(2.13)\qquad \kappa := \alpha N^2/(N^2-1) \quad\text{and}\quad r_m := (N-1)N^{-m-1}.$$

Consequently, we fix our attention to a *particular jump rate* κ and *special level probabilities*, in fact the levels are chosen according to a specific geometric law. Then, by (2.5) and (2.6),

$$(2.14)\qquad p_0 = (N-1)/N,\qquad p_\xi = N^{-2\|\xi\|},\quad \xi\neq 0.$$

In other words, from now on we consider the *continuous time random walk in* Ξ *with generator (intensity matrix)* q given in (1.9) and (1.10). By Lemma 2.7, on Ξ^*_m, $m\geq 1$, the characteristic function $\hat{p}$ of p takes on the values

$$(2.15)\qquad f_m = 1 - N^{-m} - N^{-m-1} < 1.$$

As $t\to\infty$, the transition probabilities $p(t,\cdot,\cdot)$ are *asymptotically uniformly distributed with an exponential speed* in the following sense:

Lemma 2.16. *Fix* $\zeta,\zeta'\in\Xi$. *There are positive constants* c,C *such that*

$$(2.17)\qquad \sum_{\xi\in\Xi} \left|p(t,\zeta,\xi)-p(t,\zeta',\xi)\right| \leq Ce^{-ct},\qquad t\geq 0.$$

Proof. Observe that $p(t,\zeta,\xi)$ only depends on $\|\zeta-\xi\|$ (for fixed t), and that $\|\xi-\zeta\| \neq \|\xi-\zeta'\|$ only for finitely many ξ. But for ξ,ζ,ζ' fixed, from (2.11) and (2.15) we conclude that $p(t,0,\xi-\zeta)$ and $p(t,0,\xi-\zeta')$ differ only by a sum of finitely many exponential terms of the form as written at the right hand side of (2.17). Thus we can find constants c,C as claimed. ■

With the jump rate κ from (2.13), we set

$$(2.18)\qquad \alpha_N := (N-1)\sum_{-\infty<k<+\infty} N^{-k}\exp[-\kappa(N+1)N^{-k-1}].$$

Note that this series converges. The following *asymptotics of the transition probabilities* $p(t,\zeta,0)$ will be useful.

Lemma 2.19. *For* $t>0$, *we consider* $T := T(t) \in \mathbb{R}$ *and* $\zeta := \zeta(t) \in \Xi$. *Then* $N^{T(t)}p(N^{T(t)},\zeta(t),0)$ *is bounded by* α_N *and converges to 0 or* α_N *as* $t\to\infty$, *depending on whether* $\lim_{t\to\infty}(T(t) - \|\zeta(t)\|) = -\infty$ *or* $+\infty$, *respectively.*

Proof. From (2.11) and (2.15),

$$N^T\ p(N^T,0,\zeta) = -\ [1 - \delta_0(\zeta)]\ N^{T-\|\zeta\|}\ \exp[-\kappa(N+1)N^{T-\|\zeta\|-1}]$$
$$+\ (N-1)\ \textstyle\sum_{m>\|\zeta\|}\ N^{T-m}\ \exp[-\kappa(N+1)N^{T-m-1}].$$

The second term at the r.h.s. can be written as

$$(N-1)\ \textstyle\sum_{k>\|\zeta\|-T}\ N^{-k}\ \exp[-\kappa(N+1)N^{-k-1}] \le \alpha_N.$$

On the other hand, the first term is non-positive and converges to zero as $|T - \|\zeta\|| \longrightarrow \infty$. Then the claim follows. ∎

Immediately from (2.12) and Lemma 2.19 (with $\zeta\equiv0$ and N^T replaced by t) we obtain the following asymptotics of transition probabilities:

Lemma 2.20. *With the constant* α_N *defined in* (2.18),

$$\sup_{\zeta\in\Xi}\ p(t,\zeta,0) = p(t,0,0) \sim \alpha_N/t \quad as \quad t\to\infty.$$

Consequently, $\int_1^\infty dt\ p(t,0,0) = \infty$, and we get

Lemma 2.21 **(recurrence)**. *The continuous time random walk in* Ξ *with generator* q *given in* (1.9) *and* (1.10) *is recurrent.*

At this point we will add some *tail probability estimates* we later use:

Lemma 2.22. *For* $r,t\ge0$: $\quad \sum_{\xi:\|\xi\|>r}\ p(t,0,\xi) \le t\kappa(N+1)N^{-r-1}$.

Proof. From (2.11) and (2.15), for $\xi\neq0$,

$$p(t,0,\xi) = -N^{-\|\xi\|}\exp[-t\kappa(1+N^{-1})N^{-\|\xi\|}] + (N-1)\textstyle\sum_{m>\|\xi\|}\ N^{-m}\exp[-t\kappa(1+N^{-1})N^{-m}].$$

Adding $0 \equiv N^{-\|\xi\|} - (N-1)\ \sum_{m>\|\xi\|}\ N^{-m}$ and setting $e(r) := 1-e^{-r}$, $r\ge0$, we can continue with

$$= N^{-\|\xi\|}\ e(t\kappa(1+N^{-1})N^{-\|\xi\|}) - (N-1)\ \textstyle\sum_{m:m>\|\xi\|}\ N^{-m}\ e(t\kappa(1+N^{-1})N^{-m}). \tag{2.23}$$

Estimate the negative term by 0, and use $e(r)\le r$, $r\ge0$, to arrive at

(2.24) $$p(t,0,\xi) \le t\kappa(1+N^{-1})N^{-2\|\xi\|}, \qquad t\ge 0,\ \xi\ne 0.$$

Together with (2.6) the claim follows. ∎

Our random walk Z has the following property: If $\|Z(0)\|\le s$ then $\|Z(N^s)\|$ is "of order" s as $s\to\infty$. To formulate a more precise statement, for $c,t\ge 0$ introduce the set

(2.25) $$\Xi(t,c) := \{\zeta\in\Xi;\ \big|\|\zeta\|-[t]\big| \le c\ell(t)\}, \qquad \ell(t) := (\log([t]\vee 1))/\log N.$$

Lemma 2.26 (speed of spread). *Given a constant $c>0$, there is a constant $C>0$ and an $s_0>1$ such that*

$$P_\zeta\Big(Z(N^s-N^r)\notin\Xi(s,2c)\Big) \le C[s]^{-2c} \quad \textit{whenever} \quad \|\zeta\|\le[s]+c\ell(s),\ s\ge s_0,\ s-2>r\ge-\infty$$

(with the convention $N^{-\infty}=0$).

Proof. We will split the event into two parts and proceed as follows. For $s>1$, from the uniform estimate in Lemma 2.20,

$$P_\zeta(\|Z(N^s-N^r)\| < [s]-2c\ell(s)) \le \text{const }(N^s-N^r)^{-1}\ \#\{\xi;\ \|\xi\| < [s]-2c\ell(s)\}.$$

However, $\#\{\xi;\|\xi\|\le t\}\le N^t$, $t\ge 0$, and, for all s sufficiently large, we can continue with $\le$ const $[s]^{-2c}$. In the other case,

(2.27) $$P_\zeta(\|Z(N^s-N^r)\| > [s]+2c\ell(s)) = \sum_\xi 1\{\|\xi+\zeta\| > [s]+2c\ell(s)\}\ p(N^s-N^r,0,\xi).$$

But by the definition of the addition in Ξ (coordinate wise modulo N)

(2.28) $$\|\xi+\zeta\| \le \|\xi\| \vee \|\zeta\|, \qquad \xi,\zeta\in\Xi,$$

and $\|\zeta\| \le [s]+c\ell(s)$ by assumption. Hence, ξ in the latter range of summation has to satisfy $\|\xi\| > [s]+2c\ell(s)$. That is, (2.27) can be continued with

$$\le \sum_\xi 1\{\|\xi\| > [s]+2c\ell(s)\}\ p(N^s-N^r,0,\xi) \le \text{const }[s]^{-2c},$$

where we used Lemma 2.22. This finishes the proof. ∎

2.c) Tails of the Hitting Time Distribution

Let $\tau=\tau_\zeta$ denote the *first hitting time* of 0 *after* leaving the initial state $Z(0)=\zeta\in\Xi$. An essential tool for the study of asymptotic properties of its distribution will be a "last exit from 0 decomposition" which we will present in a moment. In fact, let p^0 denote the step distribution p of (2.14) but conditioned to "proper" steps:

(2.29) $$p_0^0 := 0, \qquad p_\xi^0 := N^{-2\|\xi\|+1}, \quad \xi\neq 0.$$

Then set $P^0 := \sum_{\xi\in\Xi} p_\xi^0 P_\xi$. That is, under P^0 we assume that $Z(0)$ is distributed according to p^0, i.e. the process "starts with an immediate jump away from 0" (compare with Palm distributions in the theory of point processes). The *last exit from* 0 *decomposition* now reads as follows:

Lemma 2.30 **(last exit from 0 decomposition).** *For* $\zeta\in\Xi$, $t\geq 0$,

$$P_\zeta(\tau\leq t) = p(t,\zeta,0) - \delta_0(\zeta)e^{-t\kappa/N} + \int_0^t ds\, \left(p(s,\zeta,0) - \delta_0(\zeta)e^{-s\kappa/N}\right)\kappa P^0(\tau>t-s).$$

Sketch of Proof. Assume for the moment that $\zeta\neq 0$. Distinguish between being at time t in 0 or not. In the latter case, decompose the interval $[0,t]$ by "equidistant points" $t_1,\dots,t_m$. If this decomposition is fine enough, there must exist neighboring points t_i,t_{i+1} such that $Z(t_i)=0$, $Z(t_{i+1})\neq 0$, and Z does not return to 0 in the remaining time $t-t_{i+1}$. But $\varepsilon^{-1}p(\varepsilon,0,\xi) \xrightarrow[\varepsilon\to 0]{} \kappa p_\xi$ if $\xi\neq 0$, and at least heuristically the formula above with vanishing δ_0-terms becomes clear. If now $\zeta=0$, then additionally observe that

(2.31) $$P_0(Z(r)\neq 0 \text{ for some } r<s, \text{ and } Z(s)=0) = p(s,0,0) - e^{-s\kappa/N}, \quad s>0,$$

since $1-p_0 = 1/N$ by (2.14). This then explains why the δ_0-terms are needed in the decomposition formula. Using methods of Chung (1967), part II, §12, this sketch of proof can be made rigorous. ■

For convenience, here we introduce the *Laplace transforms*

(2.32) $$L(\lambda) := \int_0^\infty dt\, e^{-\lambda t}\, p(t,0,0), \quad L^0(\lambda) := \int_0^\infty dt\, e^{-\lambda t}\, P^0(Z(t)=0), \quad \lambda>0,$$

and formulate the following simple lemma.

Lemma 2.33. *With the constant* a_N *defined in* (2.18),

(2.34) $$L(\lambda) \sim L^0(\lambda) \sim a_N \log(1/\lambda) \quad \text{as} \quad \lambda\to 0.$$

Proof. Conditioning on the first jump time point from 0 away, as in (2.31),

(2.35) $$p(t,0,0) = e^{-t\kappa/N} + \int_0^t ds\, (\kappa/N)e^{-s\kappa/N}\, P^0(Z(t-s)=0), \quad t>0.$$

Passing to Laplace transforms this reads as

(2.36) $$L(\lambda) = (\lambda + \kappa/N)^{-1} + (\kappa/N)(\lambda + \kappa/N)^{-1}\, L^0(\lambda).$$

By Lemma 2.20, $p(t,0,0) \sim a_N/t$, hence, $\int_1^t ds\, p(s,0,0) \sim a_N \log t$ as $t\to\infty$, and a *Tauberian Theorem* (see e.g. Theorem 2 in Feller (1971), §13.5) yields

$$L(\lambda) \sim \alpha_N \log(1/\lambda) \quad \text{as} \quad \lambda \to 0.$$

But then the claim follows from (2.36). ■

Now we are ready to turn to

Proposition 2.37 **(hitting time tails).** *For the hitting time* τ *of* 0 *we have*

$$P^0(\tau > t) \sim 1/\kappa\alpha_N \log t \quad \textit{as } t \to \infty,$$

with κ *and* α_N *defined in* (2.13) *and* (2.18), *respectively.*

Proof. Fix the attention on the opposite event $\{\tau \leq t\}$. Randomizing the initial state ζ, from Lemma 2.30 we obtain the *recursion formula*

$$P^0(\tau \leq t) = P^0(Z(t)=0) + \int_0^t ds\ P^0(Z(s)=0)\ \kappa P^0(\tau > t-s). \tag{2.38}$$

Pass to Laplace transforms to get $H(\lambda) = \lambda^{-1} - L^0(\lambda) - \kappa L^0(\lambda)H(\lambda)$, where $H(\lambda) := \int_0^\infty dt\ e^{-\lambda t}\ P^0(\tau > t)$, $\lambda > 0$, and L^0 was defined in (2.32). Then

$$H(\lambda) = \lambda^{-1}(1 - \lambda L^0(\lambda))/(1 + \kappa L^0(\lambda)), \qquad \lambda > 0,$$

and from Lemma 2.33 we conclude that $H(\lambda) \sim \lambda^{-1}/\kappa\alpha_N \log(1/\lambda)$ as $\lambda \to 0$. But $P^0(\tau > t)$ is monotone in t, and by a Tauberian Theorem (see e.g. Theorem 4 in Feller (1971), §13.5) the claim follows. ■

Remark 2.39. Similarly to (2.35) and Lemma 2.33, for the distribution of the *return time* τ_0 to 0 one gets the same logarithmic tails as in the latter proposition. □

2.d) A Related Renewal Equation

In this subsection we compile some facts on a renewal equation related to the transition probability $p(\cdot,0,0)$ we need later. Write

$$g*h(t) := \int_0^t ds\ g(t-s)\ h(s), \qquad t \geq 0,$$

for the *convolution* of appropriate functions g and h.

Lemma 2.40. *Let* $A \geq 0$ *be a continuous bounded function on* $\mathbb{R}_+$ *and* $B > 0$ *a constant. Then the renewal equation*

$$h = A - B\ p(\cdot,0,0) * h \tag{2.41}$$

has a unique bounded solution h, *and this* h *is non-negative. If* A *is even a* *positive constant, then h is monotone non-increasing, and*

$$Bh(t) \sim A/\alpha_N \log t \quad \text{as } t\to\infty$$

with the constant α_N *from* (2.18).

Proof. Existence, uniqueness, non-negativity and hence boundedness of solutions follow by standard iteration arguments. Assume now that A is a positive constant. By the formulas (2.11) and (2.15),

$$p(t,0,0) = (N-1) \sum_{m>0} N^{-m} \exp[-t\kappa(N+1)N^{-m-1}], \quad t\geq 0. \tag{2.42}$$

Hence, the derivative $p'(\cdot,0,0)$ of $p(\cdot,0,0)$ is non-positive. We may differentiate the renewal equation to get

$$h' = - B\, h(t) + B \int_0^t ds\, (-p')(t-s,0,0)\, h(s), \quad t\geq 0.$$

In both terms at the r.h.s., bound h by a constant and then integrate to conclude that h' is bounded. Differentiate the equation (2.41) in the form

$$-h(t) = -A + B \int_0^t ds\, p(s,0,0)\, h(t-s), \quad t\geq 0,$$

to obtain

$$-h'(t) = Bp(t,0,0)A - B \int_0^t ds\, p(s,0,0)\, (-h')(t-s), \quad t\geq 0.$$

Since $p\geq 0$ is bounded, this is a renewal equation of the same type, and we know already it has a unique non-negative solution. Hence, $-h'$ is non-negative, consequently h is non-increasing. In order to derive the asymptotic behavior of h we use Laplace transforms. Indeed, (2.41) yields $H(\lambda) = A\lambda^{-1}/(1+BL(\lambda))$, $\lambda>0$, where H and L are the Laplace transforms of h and $p(\cdot,0,0)$, respectively. By Lemma 2.33, $L(\lambda) \sim \alpha_N \log(1/\lambda)$ as $\lambda\to 0$. Thus $H(\lambda) \sim A\lambda^{-1}/\alpha_N B\log(1/\lambda)$, and again by Tauber and the monotonicity of h the asymptotics of h follows as claimed. ∎

2.e) Asymptotic Probabilities of Hitting Times for Escaping Points

The purpose of this subsection is to show that starting at a distance $[\alpha t]$ from 0 the probability to hit 0 first after time $N^{\beta t}$, $\beta>\alpha$, converges to α/β (compare with a result of Erdös and Taylor (1960) concerning the discrete time simple random walk on $\mathbb{Z}^2$). In fact, this convergence takes place with some uniformities (recall that τ is the hitting time of 0 after leaving the initial state):

Proposition 2.43 **(limiting hitting probabilities).** *Fix constants* $0<\beta_- \le \beta_+ < +\infty$. *For* $t>0$ *let be given* $\alpha:=\alpha(t)\ge 0$, $\beta:=\beta(t)\in[\beta_-,\beta_+]$, $\rho:=\rho(t)\in[-\infty,\alpha(t)]$ *and* $\zeta:=\zeta(t)\in\Xi$. *Assume that* $\|\zeta(t)\| = \alpha(t)t + o(t)$ *as* $t\to\infty$. *Then*

$$\left|P_{\zeta(t)}\left(\tau > N^{\beta(t)t} - N^{\rho'(t)t}\right) - \alpha'(t)/\beta(t)\right| \xrightarrow[t\to\infty]{} 0 \tag{2.44}$$

where $\alpha':=\alpha\wedge\beta$ *and* $\rho':=\rho\wedge\beta$.

Proof. Roughly speaking, the main contribution to the probability of the event in (2.44) will come from last exit from **0** times included in the time interval $[N^{\alpha'(t)t},N^{\beta(t)t}]$. We shall show piece by piece that all other cases are negligible.

However, first we treat the simple and degenerate case where $N^{\beta(t)t}-N^{\rho'(t)t}$ is bounded from above along some subsequence $t'\to\infty$. Then $\rho'(t')\sim\beta(t')$ as $t'\to\infty$ since $\beta_->0$, and the second term in (2.44) tends to 1 as $t'\to\infty$, because $\rho\le\alpha$ holds. Also, here we conclude that $\|\zeta(t')\|\xrightarrow[t'\to\infty]{}\infty$ which implies that the probability term in (2.44) itself converges to 1 along $t'\to\infty$, too. Hence (2.44) is true along such subsequence $t'\to\infty$.

From now on we may assume that $N^{\beta(t)t}-N^{\rho'(t)t} \xrightarrow[t\to\infty]{} +\infty$, in particular, $\beta>\rho'=\rho$. From the last exit from **0** decomposition formula in Lemma 2.30,

$$\begin{aligned} P_\zeta(\tau \le N^{\beta t}-N^{\rho t}) &= p(N^{\beta t}-N^{\rho t},\zeta,\mathbf{0}) - \delta_{\mathbf{0}}(\zeta)\exp[-(N^{\beta t}-N^{\rho t})\kappa/N] \\ &\quad + \int_0^{N^{\beta t}-N^{\rho t}} ds\left(p(s,\zeta,\mathbf{0}) - \delta_{\mathbf{0}}(\zeta)\exp[-s\kappa/N]\right)\kappa P^{\mathbf{0}}(\tau>N^{\beta t}-N^{\rho t}-s). \end{aligned} \tag{2.45}$$

First recall that by (2.31)

$$0 \le \delta_{\mathbf{0}}(\zeta)\exp[-s\kappa/N] \le p(s,\zeta,\mathbf{0}), \qquad \zeta\in\Xi,\ s>0.$$

Hence in error estimates later on we can always neglect the δ-terms in the formula (2.45). Because of the rough bound in Lemma 2.20, the first term on the r.h.s. of (2.45) will disappear as $t\to\infty$. It remains to deal with the integral term. We will split the integral and proceed in two steps. Fix a constant $\vartheta\in(0,1)$. Of course, we may restrict the following considerations to those t satisfying $\vartheta(N^{\beta t}-N^{\rho t}) > 1$. First of all, by Lemma 2.20 we conclude that

$$\int_{\vartheta(N^{\beta t}-N^{\rho t})}^{N^{\beta t}-N^{\rho t}} ds\ p(s,\zeta,\mathbf{0})\kappa P^{\mathbf{0}}(\tau>N^{\beta t}-s) \le \text{const}\ \vartheta^{-1}(N^{\beta t}-N^{\rho t})^{-1}\int_0^{N^{\beta t}-N^{\rho t}} ds\ P^{\mathbf{0}}(\tau>s).$$

Since $N^{\beta(t)t}-N^{\rho(t)t} \xrightarrow[t\to\infty]{} \infty$ by assumption, and because the first hitting time τ of 0 is finite a.s. (by the recurrence of the random walk according to Lemma 2.21), we know that the Cesaro limit of $P^0(\tau>s)$ as $s\to\infty$ equals zero. Consequently, this part of the integral results in a negligible term.

In the second step we treat the remaining integral part from (2.45). Substituting $s=:N^{\gamma t}$, that expression can be rewritten as

$$(2.46)\qquad (t\log N)\int d\gamma\ 1\{N^{\gamma t} \le \vartheta(N^{\beta t}-N^{\rho t})\}$$
$$N^{\gamma t}\Big(p(N^{\gamma t},\zeta,0) - \delta_0(\zeta)\exp[-N^{\gamma t}\kappa/N]\Big)\ \kappa P^0(\tau>N^{\beta t}-N^{\rho t}-N^{\gamma t}).$$

Since

$$N^{\beta t}-N^{\rho t}-N^{\gamma t} \ge (1-\vartheta)(N^{\beta t}-N^{\rho t}) \xrightarrow[t\to\infty]{} +\infty$$

in that domain of integration, by the *hitting time tails* Proposition 2.37 we get

$$\kappa P^0(\tau>N^{\beta t}-N^{\rho t}-N^{\gamma t}) \sim 1/a_N \log(N^{\beta t}-N^{\rho t}-N^{\gamma t}) \sim 1/a_N\log(N^{\beta t}-N^{\rho t})$$

as $t\to\infty$, uniformly for those γ appearing in (2.46). Thus, instead of (2.46) we may investigate

$$(2.47)\qquad (t\log N)\,(\log(N^{\beta t}-N^{\rho t}))^{-1}\int d\gamma\ 1\{N^{\gamma t} \le \vartheta(N^{\beta t}-N^{\rho t})\}$$
$$a_N^{-1}\ N^{\gamma t}\Big(p(N^{\gamma t},\zeta,0) - \delta_0(\zeta)\exp[-N^{\gamma t}\kappa/N]\Big)$$

as $t\to\infty$. Take $\varepsilon>0$. Split the domain of integration into $\{\gamma \le \alpha(t)-\varepsilon\}$ and $\{\alpha(t)-\varepsilon < \gamma\}$. In the first case we apply the boundedness property formulated in Lemma 2.19 (with γt instead of $T(t)$) and use additionally $1-e^{-a}\le a$ to get

$$\le (\log(N^{\beta t}-N^{\rho t}))^{-1}\int_{-\infty}^{\alpha-\varepsilon} d\gamma\ (t\log N)\kappa(N+1)N^{\gamma t-\|\zeta\|}.$$

But $\gamma t-\|\zeta(t)\| \le -\varepsilon t + o(t)$ which is negative for all t sufficiently large. Hence, there we can integrate to obtain:

$$= \text{const}\ \ (\log(N^{\beta t}-N^{\rho t}))^{-1}\ N^{-\varepsilon t+o(t)} \xrightarrow[t\to\infty]{} 0.$$

To deal with the second case $\{\alpha(t)-\varepsilon < \gamma\}$, suppose first that the upper integration bound $(t\log N)^{-1}\log(\vartheta(N^{\beta t}-N^{\rho t})) =: h(\vartheta,t) =: h(\vartheta)$ converges to 0 along some subsequence $t'\to\infty$. Then $\beta(t')\sim\rho(t')$ as $t'\to\infty$ since $\beta_->0$, and $\alpha(t')$ will be bounded away from 0 as $t'\to\infty$. Hence $\alpha(t')-\varepsilon$ will be positive and bounded away from 0 as $t'\to\infty$, too, provided that $\varepsilon>0$ is chosen sufficiently small. Hence, the domain of integration $\{\alpha(t')-\varepsilon < \gamma \le h(\vartheta,t')\}$ will be empty for all t' sufficiently large. Consequently, along such subsequence $t'\to\infty$ the expression in (2.47) restricted to $\{\alpha(t)-\varepsilon < \gamma\}$ will disappear.

It remains to deal with the case where $h(\vartheta,t)>0$ is bounded away from 0 as $t\to\infty$, and to study

$$(2.48)\qquad (h(\vartheta))^{-1}\int_{\alpha-\varepsilon}^{h(\vartheta)} d\gamma\ a_N^{-1}\ N^{\gamma t}\left(p(N^{\gamma t},\zeta,0) - \delta_0(\zeta)\ \exp[-N^{\gamma t}\kappa/N]\right).$$

If we restrict here the domain of integration further to $\{\gamma \le \alpha(t)+\varepsilon\}$ then by using the boundedness property formulated in Lemma (2.19) we get the estimate $\le$ const ε. Letting $\varepsilon\to 0$, this leads to a negligible term. Finally, for γ in $(\alpha+\varepsilon,h(\vartheta)]$ the integrand in (2.48) converges uniformly to 1 as $t\to\infty$, namely by Lemma 2.19 since $\gamma(t)t - \|\zeta(t)\| \ge \varepsilon t + o(t) \underset{t\to\infty}{\longrightarrow} +\infty$, and because γ is bounded away from 0. Hence it remains to prove that $(h(\vartheta))^{-1}(h(\vartheta) - \alpha - \varepsilon)^+$ approaches $1 - \alpha'(t)/\beta(t)$. In fact, it suffices to study

$$(2.49)\qquad (h(\vartheta))^{-1}\ (h(\vartheta) - \alpha)^+ = 1 - 1\wedge(\alpha/h(\vartheta))$$

since the difference of both expressions is bounded in absolute value by $\le$ const ε. The proof will be finished if we show that

$$(2.50)\qquad 1\wedge(\alpha(t)/h(\vartheta,t)) - 1\wedge(\alpha(t)/\beta(t)) \underset{t\to\infty}{\longrightarrow} 0.$$

But $h(\vartheta) \sim h(1) = \beta + (t\log N)^{-1}\ \log[1 - N^{-(\beta t-\rho t)}]$ as $t\to\infty$. The latter term has its value in the interval $[-\beta_-,0)$, and we may assume that it converges along $t'\to\infty$. If its limit is 0, we are done. Otherwise, $\rho(t')\sim\beta(t')$ as $t'\to\infty$ follows, which implies that $\alpha(t')\wedge\beta(t') \sim \beta(t')$, and both terms in (2.50) will converge to 1 as $t'\to\infty$. ■

2.f) Probability Estimates for Locations at "Collision" Times

From now on, denote by $Z(\xi) = \{Z(\xi,t);t\ge 0\}$ our continuous time random walk starting in $\xi\in\Xi$. In this subsection we provide some estimates on the relative positions for a collection of four *independent* walks $(Z(\zeta^i))_{1\le i\le 4}$ needed in the next subsection to estimate collision probabilities. The *basic probability space* on which *all* our random objects are defined will be denoted by $[\Omega,\mathfrak{F},\mathcal{P}]$.

We will consider the case where the initial particles with positions $\zeta^1,\dots,\zeta^4$ may move with some restrictions. We will require that the walks considered will have a collision after a "large" time $N^{T(t)}$ and will estimate the probability that at such time N^s where two walks have the same position, the remaining walk particles will have a distance "*different* of order" s (recall (2.25)).

Lemma 2.51. *Fix* $c \geq 1$. *For* $t>0$, *let* $\zeta^1:=\zeta^1(t),\dots,\zeta^4:=\zeta^4(t) \in \Xi$ *with* $\|\zeta^4-\zeta^i\| \leq [T(t)] + c\ell(T(t))$, $i=1,2$, *where* $T(t)>0$. *Then the integrals*

$$\int_{T(t)}^{\infty} ds\; N^s\, \mathcal{P}\Big(Z(\zeta^3,N^s)-Z(\zeta^1,N^s) = 0,\quad Z(\zeta^4,N^s)-Z(\zeta^i,N^s) \notin \Xi(s,2c)\Big) \tag{2.52}$$

$i=1,2$, *converge to* 0 *as* $t\to\infty$, *provided that* $T(t) \xrightarrow[t\to\infty]{} \infty$.

Proof. Fix $Z(\zeta^1,N^s)$ for the moment. Then, by independence, from the probability expression in (2.52) we can split up $\mathcal{P}(Z(\zeta^3,N^s)-\xi=0 \,|\, Z(\zeta^1,N^s)=\xi)$, and by Lemma 2.20 estimate this term to $\leq$ const N^{-s}. Thus, the integrand in (2.52) can be estimated from above by

$$\leq \text{const}\; \mathcal{P}(Z(\zeta^4,N^s)-Z(\zeta^i,N^s) \notin \Xi(s,2c)) = \text{const}\; \mathcal{P}(\bar{Z}(\zeta^4-\zeta^i,N^s) \notin \Xi(s,2c))$$

where $\bar{Z}$ is distributed as Z except replacing the jump rate κ by 2κ. In fact, by the spatial homogeneity and symmetry of the generator q, the difference of two independent random walks each with generator q is a random walk with generator $2q$, i.e. it differs from the original walk only in the jump rate κ. By Lemma 2.26 with $r=-\infty$, for all s sufficiently large, we may continue with $\leq$ const $[s]^{-2c}$. This yields the claim since $c\geq 1$ by assumption. ■

2.g) Probability of Non-collision for Finite Systems of Independent Walks

The purpose of this subsection is to calculate the limiting probability of non-collision for a finite system of independent walks whose initial points spread out in space with time. First we introduce the following objects and notation.

Let $Z^I:=(Z(\zeta^i))_{i\in I}$ be a finite system of *independent* walks (with generator q) on the basic probability space $[\Omega,\mathfrak{F},\mathcal{P}]$, with (deterministic) starting points $\zeta^i \in \Xi$. Set $\hat{I} := \{\{i,j\};\ i,j\in I,\ i\neq j\}$. For $\{i,j\}\in\hat{I}$ let $\sigma_{i,j}\in(0,+\infty]$ denote the *first collision time* of the walks $Z(\zeta^i)$ and $Z(\zeta^j)$ *after* (!) at least one of them had left its initial state. Put $\sigma_I := \min\{\sigma_{i,j};\{i,j\}\in\hat{I}\}$ for the *first collision time* of the whole system.

Obviously, the following *decomposition formula* holds: For $0\le s\le t$, $\{i,j\}\in\hat{I}$,

(2.53) $$\mathcal{P}(\sigma_{i,j}\le N^t) = \mathcal{P}(\sigma_I=\sigma_{i,j}\le N^t) + \mathcal{P}(\sigma_I<\sigma_{i,j}\le N^t,\ \sigma_I\le N^s)$$
$$+ \sum_{(i',j')\in\hat{I}} 1\{\{i',j'\}\ne\{i,j\}\} \int_s^t \mathcal{P}(\sigma_I=\sigma_{i',j'}\in dN^r;\ \sigma_{i,j}\le N^t).$$

Write $I:=\{1,\dots,n\}$, where $n\ge 1$ is fixed. Fix also constants $c\ge 1$, $0<\beta_-\le\beta_+<+\infty$, and, for the moment, $t>0$. Consider $\alpha:=\alpha(t)\ge 0$, $\beta:=\beta(t)\in[\beta_-,\beta_+]$, $\rho:=\rho(t)\in[-\infty,\alpha(t)]$, and t-dependent starting points $\zeta^i:=\zeta^i(t)$, $i\in I$. *Assume*

(2.54) $$\zeta^i-\zeta^j \in \Xi(\alpha(t)t,c), \qquad \{i,j\}\in\hat{I}.$$

The next lemma states that on certain scales all $\binom{n}{2}$ pairs $\{i,j\}\in\hat{I}$ behave *asymptotically independent.*

Proposition 2.55 (limiting collision probability). *Under above conditions,*

(2.56) $$\left|\mathcal{P}(\sigma_I > N^{\beta(t)t}-N^{\rho'(t)t}) - (\alpha'(t)/\beta(t))^{\binom{n}{2}}\right| \underset{t\to\infty}{\longrightarrow} 0$$

where again $\alpha':=\alpha\wedge\beta$ and $\rho':=\rho\wedge\beta$.

Proof. 1°. First we want to show that we may restrict ourselves to the case $\alpha(t)\le\alpha_+$ for some bound α_+. In fact, if $\alpha(t)\underset{t\to\infty}{\longrightarrow}\infty$ then $\alpha'(t)=\beta(t)$ for all t sufficiently large. On the other hand, from (2.54) we get $\|\zeta^i-\zeta^j\|=\tilde{\alpha}_{i,j}(t)t$ with $\tilde{\alpha}_{i,j}(t)\longrightarrow\infty$ as $t\to\infty$, for each $\{i,j\}\in\hat{I}$. But

$$\mathcal{P}(\sigma_I \le N^{\beta(t)t}-N^{\rho'(t)t}) \le \sum_{(i,j)\in\hat{I}} \mathcal{P}(\sigma_{i,j} \le N^{\beta t}-N^{\rho' t}) \underset{t\to\infty}{\longrightarrow} 0$$

by Proposition 2.43. Indeed, the first collision time $\sigma_{i,j}$ of the walks $Z(\zeta^i)$ and $Z(\zeta^j)$ coincides in distribution with the hitting time τ_ξ of 0 of a single random walk with generator $2q$ starting in $\xi:=\zeta^i-\zeta^j$, and $\tilde{\alpha}'_{i,j}(t)=\beta(t)$ for t sufficiently large.

2°. To show that it suffices to consider the case $\rho\equiv-\infty$ look at:

$$\mathcal{P}(\sigma_I > N^{\beta t}-N^{\rho' t}) - \mathcal{P}(\sigma_I > N^{\beta t}) = \mathcal{P}(\sigma_I \in (N^{\beta t}-N^{\rho' t},N^{\beta t}])$$
$$\le \sum_{(i,j)\in\hat{I}} \mathcal{P}(\sigma_{i,j} \in (N^{\beta t}-N^{\rho' t},N^{\beta t}]).$$

Because of $\alpha(t)\le\alpha_+$, from (2.54) we conclude $\|\zeta^i-\zeta^j\| = \alpha(t)t + o(t)$ as $t\to\infty$, $\{i,j\}\in\hat{I}$. Applying twice Proposition 2.43 we recognize that each of those finitely many summands tends to 0 as $t\to\infty$. Similarly, without loss of generality we may assume that there is a constant $\alpha_->0$ such that $\alpha(t)\ge\alpha_-$, $t>0$, since the event $\sigma_I > N^{\beta t}$ implies the existence of a pair $\{i,j\} \in \hat{I}$ such that $\sigma_{i,j} > N^{\beta t}$, and $\alpha(t) \xrightarrow[t\to\infty]{} 0$ leads to a vanishing limiting probability.

3°. Fix the attention on the opposite event $\{\sigma_I \le N^{\beta(t)t}\}$. Assume for the moment that starting from the decomposition formula (2.53) (with t and $\beta(t)$t instead of s and t) we can derive the following *integral equation:*

$$(2.56) \qquad (1 - \alpha'/\beta) = \mathcal{P}(\sigma_I = \sigma_{i,j} \le N^{\beta t}) + \varepsilon(t,\alpha,\beta,Z^I(0)) + \beta^{-1} \sum_{(i',j')\neq(i,j)} \int_{\alpha'}^{\beta} d\gamma\ \mathcal{P}(\sigma_I = \sigma_{i',j'} \le N^{\gamma t}), \qquad \{i,j\}\in\hat{I},\ t>0.$$

Here and below ε always denotes some functions depending (among other things) on t and converging to 0 as $t\to\infty$. Summing over the $\binom{n}{2}$ pairs $\{i,j\}\in\hat{I}$ yields

$$\binom{n}{2}(1 - \alpha'/\beta) = \mathcal{P}(\sigma_I \le N^{\beta t}) + \varepsilon(t,\alpha,\beta,Z^I(0)) + \beta^{-1}\left(\binom{n}{2}-1\right) \int_{\alpha'}^{\beta} d\gamma\ \mathcal{P}(\sigma_I \le N^{\gamma t}).$$

Then we can use the following asymptotic result on an "abstract" *integral equation* (which follows from a simple contraction argument, see [8], Lemma 2): If for a constant $K\ge 1$ the bounded measurable function q satisfies

$$K(1 - \alpha'(t)/\beta(t)) = q(t,\beta(t)) + \frac{K-1}{\beta(t)} \int_{\alpha'(t)}^{\beta(t)} d\gamma\ q(t,\gamma) + \varepsilon(t),$$

where $0<\alpha_0\le\alpha'(t)\le\beta(t)\le\beta_0<+\infty$, $t>0$, and $\varepsilon(t)\to 0$ as $t\to\infty$ according to our convention, then

$$\left| q(t,\beta(t)) - [1 - (\alpha'(t)/\beta(t))^K] \right| \longrightarrow 0 \quad \text{as } t\to\infty.$$

It therefore remains to prove (2.56) by starting from (2.53).

4°. We discuss the single terms in (2.53) separately (with $[s,t] = [\alpha'(t)t, \beta(t)t]$). For the left hand side, inserting the *limiting hitting probabilities* according to our Proposition 2.43 (which are *independent* of the jump rate)

$$\mathcal{P}(\sigma_{i,j} \le N^{\beta t}) = (1 - \alpha'/\beta) + \varepsilon(t,\alpha,\beta,Z^I(0)) \quad \text{as } t\to\infty.$$

By the same proposition, $\mathcal{P}(\sigma_{i',j'} \le N^{\alpha' t}) = \varepsilon(t,\alpha,\beta,Z^I(0))$ as $t\to\infty$, which we apply to the second term at the r.h.s. of (2.53) to recognize that it is a negligible term. The integral term in (2.53) we will split into two pieces. First the integral over the event that at the time N^r of the first collision of $Z(\zeta^{i'})$ and $Z(\zeta^{j'})$ the walks $Z(\zeta^i)$ and $Z(\zeta^j)$ have a relative position outside of $\Xi(r,2c)$ (recall (2.25)), and second the remaining part. Then for the first part we get the upper estimate

$$\le \sum_{(i',j')\ne(i,j)} \int_{\alpha' t}^{\beta t} dN^r \; \mathcal{P}\Big(Z(\zeta^{i'},N^r)=Z(\zeta^{j'},N^r),\; Z(\zeta^{i},N^r)-Z(\zeta^{j},N^r) \notin \Xi(r,2c)\Big).$$

Since the latter integral disappears if $\alpha'(t)\le\alpha(t)$, the domain of integration can be replaced by $[[\alpha(t)t],+\infty)$. Apply Lemma 2.51 with $T(t)=[\alpha(t)t]$ to each of these finitely many summands to see that this results into a term $\varepsilon(t,Z^I(0))$. It remains to consider those walk realizations $Z(\zeta^i)$ and $Z(\zeta^j)$ which have at time N^r positions $\zeta,\zeta'\in\Xi$ with $\zeta-\zeta'\in\Xi(r,2c)$, which implies that $\|\zeta-\zeta'\| = r + o(r)$ as $r\to\infty$. They have to collide in the remaining time $N^{\beta(t)t}-N^r$. By Proposition 2.43 with setting there $\rho(t)=\alpha(t)=r(t)/t$, the latter collision probability equals $(1 - r/\beta t) + \varepsilon(t,r/t,\beta,\zeta,\zeta')$, uniformly in those r,ζ,ζ'. With the strong Markov property, we can summarize all these asymptotic relations as $t\to\infty$ as follows:

$$(1 - \alpha'/\beta) \;=\; \varepsilon(t,\alpha,\beta,Z^I(0)) + \mathcal{P}(\sigma_I=\sigma_{i,j}\le N^{\beta t})$$
$$+ \sum_{(i',j')\ne(i,j)} \int_{\alpha' t}^{\beta t} \mathcal{P}(\sigma_I=\sigma_{i',j'}\in dN^r)\;(1 - r/\beta t).$$

Replace the integrand by $(1/\beta t)\int_s^{\beta t} dr$, change the order of integration, substitute $r=\gamma t$, and apply again Proposition 2.43 to rewrite the integral as

$$= \beta^{-1} \int_{\alpha'}^{\beta} d\gamma \; \mathcal{P}(\sigma_I=\sigma_{i',j'}\le N^{\gamma t}) + \varepsilon(t,\alpha,\zeta^{i'},\zeta^{j'}).$$

This gives (2.56) and finishes the proof. ■

2.h) On Moments of Interacting Diffusion Systems

We return to our interacting diffusion system $\mathfrak{X}$ as in Definition 1.6, starting with a deterministic initial state $\mathbf{z}\in[0,1]^{\Xi}$. Write $\mathbb{P}^{g}_{\mathbf{z}}$ instead of $\mathbb{P}^{g}_{\delta_{\mathbf{z}}}$. We want to show how *independent* random walks in Ξ can be used to describe moments of $\mathfrak{X}$, a technique we shall need later.

But first we mention a *technical point.* In comparison arguments in the Sections 4-6 we need a notion of Fisher-Wright diffusions on a *sub*interval of $[0,1]$. Therefore at some places we will replace the assumption of (strict) positivity of g on the whole open interval $(0,1)$, imposed in (1.2) and Definition 1.6. Let $\mathcal{G}$ denote the set of all those functions g satisfying

(2.57) $\quad g\colon [0,1]\mapsto\mathbb{R}_+$ is Lipschitz continuous, $g>0$ only on an open subinterval.

Note that the Definition 1.6 still makes sense (see [27]) if we replace $\mathcal{G}^0$ by

$\mathcal{G}\supset\mathcal{G}^0$; we will apply all notation introduced also in the more general situation.

To describe the moments we need some further objects. Fix $n\geq 1$, set $I:=\{1,\dots n\}$, and write $\bar\zeta:=[\zeta(1),\dots,\zeta(n)]\in\Xi^I$. Consider a system of n *independent* random walks in Ξ each with generator q, denote the generator of this system by $\mathfrak{A}^I$, the corresponding semi-group by $\mathcal{S}^I$, and the transition probabilities by p^I. In a direct generalization of Lemma 1 in Cox and Greven (1991b), where $n\leq 2$, we get:

Lemma 2.58. **(moment equations).** *Let $g\in\mathcal{G}$. Then the interacting diffusion system $\mathfrak{X}=[\mathfrak{x},\mathbb{P}^g_\mu,\mu\in\mathfrak{P}]$ satisfies the following moment equations. For all $z\in[0,1]^\Xi$, vectors $\bar\zeta=[\zeta(1),\dots,\zeta(n)]\in\Xi^I$, and $t\geq 0$,*

$$\mathbb{E}^g_{\mathbf{z}}\prod_{i\in I}\mathfrak{x}_{\zeta(i)}(t) = \sum_{\bar\xi\in\Xi^I} p^I(t,\bar\zeta,\bar\xi)\prod_{m\in I} z_{\xi(m)}$$
$$+\int_0^t ds\sum_{\bar\xi\in\Xi^I} p^I(t-s,\bar\zeta,\bar\xi)\Big(\sum_{\{i,j\}\in\hat I}\delta_{\xi(i),\xi(j)}\,\mathbb{E}^g_{\mathbf{z}} g(\mathfrak{x}_{\xi(i)}(s))\prod_{m\in I\setminus\{i,j\}}\mathfrak{x}_{\xi(m)}(s)\Big).$$

(with $\hat I$ the set of all pairs of I, as in the previous subsection, and δ the Kronecker symbol). In particular, the first moments are explicitly given by

(2.59) $$\mathbb{E}^g_{\mathbf{z}}\mathfrak{x}_\zeta(t) = \sum_\xi p(t,\zeta,\xi)z_\xi =: a_{\mathbf{z},\zeta}(t),\qquad z\in[0,1]^\Xi,\ \zeta\in\Xi,\ t\geq 0,$$

and do not depend on the diffusion coefficient $g\in\mathcal{G}$.

Proof. Let $U=U^g$ denote the (Feller) *semigroup* of $\mathfrak{X}$ acting on $\mathbf{C}[[0,1]^\Xi]$. The corresponding *generator* $\mathfrak{G}=\mathfrak{G}^g$ is the closure of the following operator acting on functions h on $[0,1]^\Xi$ which depend only on finitely many components in a twice continuously differentiable way (compare Shiga and Shimizu (1980)):

(2.60) $$(\mathfrak{G}h)(\mathbf{x}) = \Big(\sum_{\zeta,\xi} q_{\zeta,\xi}(\mathbf{x}_\xi-\mathbf{x}_\zeta)(\partial/\partial\mathbf{x}_\zeta) + 2^{-1}\sum_\zeta g(\mathbf{x}_\zeta)(\partial/\partial\mathbf{x}_\zeta)^2\Big)h(\mathbf{x}),\quad \mathbf{x}\in[0,1]^\Xi.$$

Fix $\bar\zeta=[\zeta(1),\dots,\zeta(n)]\in\Xi^I$ for the moment and define the following function $h_{\bar\zeta}(\mathbf{x}) := \prod_{i\in I}\mathbf{x}_{\zeta(i)}$, $\mathbf{x}\in[0,1]^\Xi$. Apply (2.60) to this function to get

$$(\mathfrak{G}h_{\bar\zeta})(\mathbf{x}) = \sum_{i\in I}\sum_{\xi\in\Xi} q_{\zeta(i),\xi}\,(\mathbf{x}_\xi-\mathbf{x}_{\zeta(i)})\prod_{j\neq i}\mathbf{x}_{\zeta(j)}$$
$$+\sum_{\{i,j\}\in\hat I}\delta_{\zeta(i),\zeta(j)}\,g(\mathbf{x}_{\zeta(i)})\prod_{m\in I\setminus\{i,j\}}\mathbf{x}_{\zeta(m)},\qquad \mathbf{x}\in[0,1]^\Xi.$$

Hence, from the U^g-semigroup property related to the Markov process $\mathfrak{X}$

$$\frac{d}{dt}\mathbb{E}^g_{\mathbf{z}}\prod_i\mathfrak{x}_{\zeta(i)}(t) = \frac{d}{dt}\mathbb{E}^g_{\mathbf{z}}h_{\bar\zeta}(\mathfrak{x}(t)) = \mathbb{E}^g_{\mathbf{z}}\mathfrak{G}h_{\bar\zeta}(\mathfrak{x}(t))$$
$$= \sum_{i\in I}\sum_{\xi\in\Xi} q_{\zeta(i),\xi}\,\mathbb{E}^g_{\mathbf{z}}\Big[\mathfrak{x}_\xi(t)\prod_{j\neq i}\mathfrak{x}_{\zeta(j)}(t) - \prod_k\mathfrak{x}_{\zeta(k)}(t)\Big]$$

$$+ \sum_{(i,j)\in\hat{I}} \delta_{\zeta(i),\zeta(j)} \mathbb{E}^{g}_{\mathbf{z}} g(\mathfrak{x}_{\zeta(i)}(t)) \prod_{m\in I\setminus(i,j)} \mathfrak{x}_{\zeta(m)}(t), \qquad z\in[0,1]^{\Xi},\ t>0.$$

Fix $z\in[0,1]^{\Xi}$ and abbreviate

$$v_t(\bar\zeta) := \sum_{(i,j)\in\hat{I}} \delta_{\zeta(i),\zeta(j)} \mathbb{E}^{g}_{\mathbf{z}} g(\mathfrak{x}_{\zeta(i)}(t)) \prod_{m\in I\setminus(i,j)} \mathfrak{x}_{\zeta(m)}(t).$$

$$u_t(\bar\zeta) := \mathbb{E}^{g}_{\mathbf{z}} \prod_{k\in I} \mathfrak{x}_{\zeta(k)}(t), \qquad t\geq 0, \quad \bar\zeta=[\zeta(1),\dots,\zeta(n)]\in\Xi^{I}.$$

Then the latter system of differential equations can be rephrased by using the generator $\mathfrak{A}^{I}$ of n independent walks introduced before the lemma:

$$\frac{d}{dt}u_t = \mathfrak{A}^{I} u_t + v_t.$$

Its solution is

$$u_t = \mathcal{S}^{I}_{t} u_0 + \int_0^t ds\, \mathcal{S}^{I}_{t-s} v_s, \qquad t\geq 0;$$

see Theorem I.2.15 in Liggett (1985). But this is nothing else than the claim of the lemma. ∎

Consider now the *Fisher-Wright* case $g=\mathfrak{b}\ell$ of (1.3). By an abuse of notation, we will often replace the upper index $g=\mathfrak{b}\ell$ by $\mathfrak{b}$. Start with the *constant initial state* $\bar\theta$, where $\bar\theta_{\xi}\equiv\theta\in(0,1)$. The speed at which the process moves to the boundary points is described in the following

Lemma 2.61. *In the case of an interacting Fisher-Wright system with constant initial state* $\bar\theta$,

$$\mathbb{E}^{\mathfrak{b}}_{\bar\theta}\ell(\mathfrak{x}_{\zeta}(t)) \sim 2\ell(\theta)/\alpha_N \mathfrak{b}\log t \quad \textit{as } t\to\infty, \quad \zeta\in\Xi,$$

where α_N *is the constant defined in* (2.18), *and* $\mathfrak{b}$ *is the diffusion constant.*

Proof. Set $\zeta=0$ by the shift-invariance of $\mathfrak{X}$. From Lemma 2.58, by the constancy in ξ of $\mathbb{E}^{\mathfrak{b}}_{\mathbf{z}}\ell(\mathfrak{x}_{\xi}(s))$, and the symmetry of the random walk, we have

$$\mathbb{E}^{\mathfrak{b}}_{\bar\theta}\ell(\mathfrak{x}_0(\cdot)) = \ell(\theta) - \mathfrak{b}\ p(2\cdot,0,0) * \mathbb{E}^{\mathfrak{b}}_{\bar\theta}\ell(\mathfrak{x}_0(\cdot)).$$

Thus, Lemma 2.40 yields the claim. ∎

3. Preparations: Coalescing Random Walks on the Hierarchical Group

This section provides the essential results on coalescing random walks on Ξ which will be needed to verify the Theorems 1-4 in the Fisher-Wright case via the duality relation. This will be formulated and proved in the Propositions 3.13 and 3.28 in the Subsections 3.d) and e) below. We start by first preparing the necessary tools in 3.a) and c):

3.a) The Dual Process η

The purpose of this subsection is to introduce a *dual process* η to the system of interacting Fisher-Wright diffusions on the interval $[0,1]$ and derive a first result. The dual process will be an interacting system with *finitely many particles,* namely a system of coalescing random walks with delay. Roughly speaking, particles move independently as random walks of the type in the previous section, except that particles at the same position *may react,* that is, each pair of particles coalesces independently into a single particle at the constant rate $\mathfrak{b}$ (which is the Fisher-Wright diffusion constant of (1.3)).

Let Φ denote the countable set of all those elements $\varphi=\{\varphi_\xi;\xi\in\Xi\}$ in $\mathbb{Z}_+^\Xi$ which have only finitely many coordinates φ_ξ different from 0. We interpret φ as a finite particle system in Ξ, where φ_ξ counts the number of particles at site ξ. The one-particle system consisting of a single particle situated at $\xi\in\Xi$ is denoted by δ^ξ. The addition in Φ is defined component wise. Write $\|\varphi\|$ for the *total number* $\sum_{\xi\in\Xi}\varphi_\xi$ of particles in $\varphi\in\Phi$. For fixed $\mathfrak{b}>0$, set

(3.1)
$$Q_{\varphi,\psi} := \begin{cases} \varphi_\zeta\, q_{\zeta,\xi} & \text{if } \psi=\varphi-\delta^\zeta+\delta^\xi, \quad \xi\neq\zeta \\ \mathfrak{b}\varphi_\zeta(\varphi_\zeta-1)/2 & \text{if } \psi=\varphi-\delta^\zeta \\ 0 & \text{otherwise} \end{cases}, \qquad \varphi,\psi\in\Phi,\ \varphi\neq\psi,$$
$$Q_{\varphi,\varphi} := -\sum_{\psi\neq\varphi} Q_{\varphi,\psi}, \qquad \varphi\in\Phi.$$

We obtain a generator Q of a continuous time Markov chain with trajectories in $\mathbf{D}[\mathbb{R}_+,\Phi]$. This is our *coalescing random walk* $\eta := (\{\eta_\xi(t);t\geq0\})_{\xi\in\Xi}$ *with delay* which has random walk generator q (defined in (1.9) and (1.10)) and *coalescing rate* $\mathfrak{b}$. The distribution of η with starting population $\varphi\in\Phi$ is denoted by $\mathbf{P}_\varphi$. In this case we also write η^φ instead of η.

To explain the crucial relation between the stochastic process η and the system of interacting Fisher-Wright diffusions $\mathfrak{X}$, we introduce the following notation. For $\mathbf{z}\in[0,1]^\Xi$ and $\varphi\in\Phi$ we write $\mathbf{z}^\varphi := \prod_{\xi\in\Xi}(\mathbf{z}_\xi)^{\varphi_\xi}$ where we set $0^0=1$. Note that this definition of $\mathbf{z}^\varphi$ makes sense since φ has only finitely many coordinates different from 0. Applying the generator $\mathfrak{G}=\mathfrak{G}^{\mathfrak{b}}$ defined in (2.60) to the function $h_\varphi(\mathbf{z}) := \mathbf{z}^\varphi$, $\mathbf{z}\in[0,1]^\Xi$, yields

$$\mathfrak{G}^{b} z^{\varphi} = \sum_{\psi\in\Phi} Q_{\varphi,\psi} z^{\psi}, \qquad z\in[0,1]^{\Xi},\ \varphi\in\Phi.$$

From this one easily derives (cf. Shiga (1980)) the following *duality relation* between the systems of interacting Fisher-Wright diffusions $\mathfrak{X}$ and coalescing random walks η with delay, which is extremely powerful.

Lemma 3.2 (duality). *For* $z\in[0,1]^{\Xi}$, $\varphi\in\Phi$, *and* $t\geq 0$: $\quad \mathbb{E}_{z}^{b}(\mathfrak{X}(t))^{\varphi} = \mathbf{E}_{\varphi} z^{\eta(t)}$.

Since for fixed t all moments $\mathbb{E}_{z}^{b}(\mathfrak{X}(t))^{\varphi}$, $\varphi\in\Phi$, determine uniquely the distribution of $\mathfrak{X}(t)=(\mathfrak{x}_{\xi}(t))_{\xi\in\Xi}$ with respect to $\mathbb{P}_{z}^{b}$, by duality it suffices to study the *generating functions* $\mathbf{E}_{\varphi} z^{\eta(t)}$, $z\in[0,1]^{\Xi}$, of $\eta(t)$ with respect to $\mathbf{P}_{\varphi}$ which is a considerably simpler task. In this sense, the infinite system of interacting Fisher-Wright diffusions is replaced by a collection of finite systems of coalescing random walks.

3.b) An Application of Duality

Occasionally it is useful to read the duality relation in the opposite direction, namely to verify a property of coalescing walks with delay. For instance, later on we need to know the behavior of

$$q_{2,2}(t) := \mathcal{P}(\|\eta(t)\|=2\,|\,\eta(0)=2\delta^{0}) \quad \text{as} \quad t\to\infty,$$

that is the conditional probability that the random walk with delay η *did not coalesce by time* t given the system started off with exactly two particles at a common location.

Lemma 3.3. *With the constant* α_N *from* (2.18),

(3.4) $$q_{2,2}(t) \sim 2/\alpha_N b \log t \quad \textit{as} \quad t\to\infty.$$

Proof. By the first moment formula (2.59) and the duality Lemma 3.2,

$$\mathbb{E}_{\theta}^{b} f(\mathfrak{x}_0(t)) = \theta - \mathbf{E}_{2\delta_0}\theta^{\|\eta(t)\|} = q_{2,2}(t)\,(\theta-\theta^2), \qquad \theta\in(0,1),\ t\geq 0.$$

Finish the proof with Lemma 2.61. ■

3.c) The Approximate Dual Process $\tilde{\eta}$

For our purpose it is useful to still simplify the right hand side of the duality relation (Lemma 3.2) by introducing a process $\tilde{\eta}$ which is for our purpose *approximately* equivalent to η as $t\to\infty$. The point is that the system η

of coalescing random walks with delay introduced above has the property that particles may or may not coalesce when they collide. Instead of that, in defining $\tilde{\eta}$ we will now require, that each pair of colliding particles coalesces *instantaneously* into a single particle.

More precisely let $\tilde{\Phi}$ denote the set of all those $\psi\in\Phi$ which satisfy $\psi_\xi\leq 1$, $\xi\in\Xi$. Define

$$\tilde{Q}_{\varphi,\psi} := \begin{cases} q_{\zeta,\xi} & \text{if } \psi = \varphi - \delta^\zeta + (1-\varphi(\xi))\delta^\xi, \quad \xi\neq\zeta \\ 0 & \text{otherwise} \end{cases}, \qquad \varphi,\psi\in\tilde{\Phi},\ \varphi\neq\psi,$$

$$\tilde{Q}_{\varphi,\varphi} := -\sum\nolimits_{\psi\neq\varphi} \tilde{Q}_{\varphi,\psi}, \qquad \varphi\in\tilde{\Phi}.$$

We obtain the generator $\tilde{Q}$ of the desired continuous time Markov chain $\tilde{\eta}$ with sample paths in $\mathbf{D}[\mathbb{R}_+,\tilde{\Phi}]$. The distribution of this *coalescing random walk* $\tilde{\eta}$ with starting population $\varphi\in\tilde{\Phi}$ is denoted by $\tilde{\mathbf{P}}_\varphi$; here again we write $\tilde{\eta}=\tilde{\eta}^\varphi$.

The fact we need is now that η and $\tilde{\eta}$ are in a sense approximately equivalent as $t\to\infty$, even when the initial states are time-dependent. Roughly speaking, if two *independent* copies of our walk meet once, by their recurrence they will meet infinitely often; and if they get each time the chance to coalesce with the fixed positive rate $\mathfrak{b}$ (as in the coalescing walk with delay), they will finally coalesce.

To obtain a more formal statement we need some notation. Associate with each $\varphi\in\Phi$ the *truncated* element $\varphi^*\in\tilde{\Phi}$ defined by $\varphi^*_\xi := \varphi_\xi\wedge 1$, $\xi\in\Xi$. On the other hand, if $Z^I=\{Z(\xi(i));i\in I\}$ is a finite system of *independent* walks with starting points $\xi(i)\in\Xi$, $i\in I$, introduced in the beginning of Subsection 2.g), then formally associate the Φ-valued Markov process $\mathcal{Z}^\varphi := \{\mathcal{Z}^\varphi(t);t\geq 0\}$ defined by $\mathcal{Z}^\varphi(t) := \sum_{i\in I} \delta^{Z(\xi(i),t)}$, $t\geq 0$, where $\varphi=\delta^{\xi(1)}+\ldots+\delta^{\xi(n)}$. Actually, we will identify Z^I and $\mathcal{Z}^\varphi$. From now on we will use the following concept of *coupling.*

Convention 3.5 **(coupling).** Choose our basic probability space $[\Omega,\mathfrak{F},\mathcal{P}]$ in such a way that it supports all three Markov families $[\mathcal{Z}^\varphi,\varphi\in\Phi]$, $[\eta^\psi,\psi\in\Phi]$, $[\tilde{\eta}^\chi,\chi\in\tilde{\Phi}]$ and that they satisfy $\mathcal{Z}^\varphi(s)\geq\eta^\psi(s)\geq\tilde{\eta}^\chi(s)$ for all $s>0$, whenever $\varphi\geq\psi\geq\chi$, $\varphi,\psi\in\Phi$, $\chi\in\tilde{\Phi}$. ∎

Now we are in a position to formulate the following *approximation* result.

Proposition 3.6 **(approximation).** *Fix* $n\ge 1$. *For* $t>1$, *let* $\varphi(t)=\delta^{\zeta(1,t)}+\dots+\delta^{\zeta(n,t)}\in\Phi$ *and assume*

(3.7) $$\|\zeta(i,t)-\zeta(j,t)\| = \alpha_{i,j}t + o(t) \quad as \quad t\to\infty, \quad i\ne j,$$

for some constants $\alpha_{i,j}\ge 0$. *Then*

(3.8) $$\mathcal{P}\left(\eta^{\varphi(t)}(N^t) = \tilde{\eta}^{\varphi^*(t)}(N^t)\right) \xrightarrow[t\to\infty]{} 1.$$

Proof. Let E_n denote the event in (3.8). Trivially, the claim holds for $n=1$; suppose that it is true for some $n-1\ge 1$. First we consider the case when $\alpha_{i,j}\ge 1$ for all $i\ne j$. Then, by (3.7) we can assume without loss of generality that $\varphi=\varphi^*$. Continue with $\mathcal{P}(\complement E_n) \le \sum_{i\ne j} \mathcal{P}(\sigma_{i,j}\le N^t) \xrightarrow[t\to\infty]{} 0$ by Proposition 2.43.

Now we come to the main case if there exists a pair $i\ne j$ with $\alpha_{i,j}<1$. Fix such $[i,j]$. For $M>1$, let $E^{M,t}_{i,j}$ denote the event that $Z(\zeta(i,t))$ and $Z(\zeta(j,t))$ meet at least M-times before time N^t and coalesce during one of these meetings. Moreover, define E^{i}_{n-1} as E_n but with $\varphi(t) - \delta^{\zeta(i,t)}$ in place of $\varphi(t)$. Then we get $\mathcal{P}(E_n) \ge \mathcal{P}(E^{i}_{n-1}\cap E^{M,t}_{i,j}) \ge \mathcal{P}(E^{i}_{n-1}) - \mathcal{P}(\complement E^{M,t}_{i,j})$. However, by Proposition 2.43, and by the recurrence according to Lemma 2.21, $\mathcal{P}(E^{M,t}_{i,j})$ converges to 1 by first letting $t\to\infty$ and then $M\to\infty$. Therefore $\liminf_{t\to\infty} \mathcal{P}(E_n) \ge \liminf_{t\to\infty} \mathcal{P}(E^{i}_{n-1})$. Apply the induction hypothesis to complete the proof. ■

3.d) Longterm Behavior of $\tilde{\eta}$ in the Case of Time-dependent Initial Points

The next task is to formulate a scaling limit proposition about the coalescing random walk $\tilde{\eta}$; compare with the corresponding property on the two dimensional lattice in Cox and Griffeath (1986), Theorem 3.

First we want to introduce the expressions for the limit probabilities. Consider the following homogeneous linear *system of integral equations:*

(3.9) $$p_{n+1,k}(\alpha) = \int_\alpha^1 d\gamma \, \binom{n+1}{2}\alpha^{\binom{n+1}{2}}\gamma^{-\binom{n+1}{2}-1} \, p_{n,k}(\gamma), \qquad 1\le k\le n, \; 0\le\alpha\le 1,$$

(3.10) $$p_{n,n}(\alpha) = \alpha^{\binom{n}{2}}, \qquad n\ge 1, \; 0\le\alpha\le 1,$$

(recall $0^0=1$). It is easy to see that this system of (continuous) functions defined on $[0,1]$ is *uniquely* solvable, and that each $p_{n,(\cdot)}(\alpha)$, $n\ge 1$, $0\le\alpha\le 1$, is a *probability* law on $\{1,\dots,n\}$ which will describe the limit probability for the number of particles living at time N^t $(t\to\infty)$ when n initial particles escape from each other with speed α. Note that

$$(3.11)\qquad p_{n+1,k}(\alpha/\beta) = \int_\alpha^\beta d_\gamma\left(1-(\alpha/\gamma)^{\binom{n+1}{2}}\right) p_{n,k}(\gamma/\beta), \qquad 1\le k\le n,\ 0\le\alpha\le\beta,\ \beta>0,$$

(with γ as the integration variable). Hence, by Proposition 2.55, we can interpret this as a *"limiting decomposition"* with respect to the first collision at time $N^{\gamma t}$ of $n+1$ initial particles (see also (2.53)).

To formulate the relevant assumptions, for $n\ge 1$, and $c,r\ge 0$, write

$$(3.12)\qquad \tilde{\Phi}(n,r,c) := \{\delta^{\zeta(1)}+\ldots+\delta^{\zeta(n)}\in\tilde{\Phi};\ \ \zeta(i)-\zeta(j)\in\Xi(r,c)\ \text{ for } i\ne j\}$$

for the set of all n particle configurations with particles' distance $[r]$, except a logarithmic error (recall (2.25)). Note that $\tilde{\Phi}(n,r,c)$ is empty if $r\le 1$ and $n\ge 2$.

Proposition 3.13 **(scaling limit).** *Fix* $n\ge 1$ *and constants* $c\ge 1$, $0<\beta_-\le\beta_+<+\infty$. *For* $t>0$, *consider* $\alpha:=\alpha(t)\ge 0$, $\beta:=\beta(t)\in[\beta_-,\beta_+]$, $\rho:=\rho(t)\in[-\infty,\alpha(t)]$, *and* $\varphi:=\varphi(t)\in\tilde{\Phi}(n,\alpha(t)t,c)$. *Then*

$$(3.14)\qquad \lim_{t\to\infty}\left|\tilde{\mathbf{P}}_{\varphi(t)}\left(\|\tilde{\eta}(N^{\beta(t)t}-N^{\rho'(t)t})\|=k\right) - p_{n,k}(\alpha'(t)/\beta(t))\right| = 0, \qquad 1\le k\le n,$$

where again $\alpha' := \alpha\wedge\beta$ *and* $\rho' := \rho\wedge\beta$.

Of course, the case $0<\alpha(t)\equiv\alpha<\beta\equiv\beta(t)$, $\rho(t)\equiv-\infty$ is of a particular interest. On the other hand, allowing a t-dependence of α and β, and adding $\rho(t)$, the relation (3.14) implies a *uniformity* in the convergence. This uniformity will be crucial in some induction arguments later on.

Remark 3.15. As pointed out in [8], by differentiation and a change of variable, (3.9) can be transformed into the *backward equations* for the transition probabilities of a *pure non-linear death process* on $\{1,2,\ldots\}$ which jumps from n to $n-1$ at rate $\binom{n}{2}$, $n\ge 2$. Calculating eigenvalues and eigenfunctions, and using the spectral representation of solutions, one can actually show (see Tavaré (1984), Appendix I) that

$$(3.16)\qquad p_{n,k}(\gamma) = \sum_{i=k}^{n} \frac{(-1)^{i+k}\,(2i-1)\,(i+k-2)!\,\binom{n}{i}}{k!\,(k-1)!\,(i-k)!\,\binom{n+i-1}{i}}\,\gamma^{\binom{i}{2}}, \qquad 1\le k\le n,\ 0\le\gamma\le 1. \qquad \square$$

Proof of Proposition 3.13. 1°. As in the second step of proof of Prop. 2.55, it suffices to consider the case $\rho(t)\equiv-\infty$. In fact, the first term in the assertion (3.14) differs from the one with $\rho(t) \equiv -\infty$ at most by

$$\le 2\, \mathcal{P}\left(\tilde{\eta}^{\varphi} \text{ has some collision in } (N^{\beta t}-N^{\rho' t}, N^{\beta t})\right).$$

But then there must exist a pair $i \neq j$ such that the *first* collision $\sigma_{i,j}$ occurs in $(N^{\beta t}-N^{\rho' t}, N^{\beta t})$. Apply twice Proposition 2.55 (with n=2) to see that this gets a vanishing probability as $t \to \infty$. Also, as in 1° and 2° of proof of our Proposition 2.55 we may assume that $\alpha(t)$ is bounded away from *0* and infinity. Note that in this case $\tilde{\Phi}(n,\alpha(t)t,c)$ is non-empty for t sufficiently large.

2°. Fix the *final* number of particles $k \ge 1$. The proof will be by induction on the number $n \ge k$ of *initial* particles. The initial step of *induction* $n=k$ is provided by (a special case of) the *limiting collision probability* Proposition 2.55. Assume that the claim is true for some $n-1 \ge k$ and all $c \ge 1$. According to (3.9) it suffices to prove that

$$\mathcal{P}(\|\tilde{\eta}^{\varphi}(N^{\beta t})\|=k) \approx \int_{\alpha'}^{\beta} d\gamma \, \binom{n}{2} \alpha'^{\binom{n}{2}} \gamma^{-\binom{n}{2}-1} \, p_{n-1,k}(\gamma/\beta) \quad \text{as } t\to\infty, \tag{3.17}$$

where we write $\approx$ for equality except an additive error term $\varepsilon(t,\alpha,\beta,\varphi)$.

Observe that by Proposition 2.55, besides an error term as $t \to \infty$, the first collision time σ_{φ} of the whole system starting in $\varphi(t)$ is not smaller than $N^{\alpha'(t)t}$. Hence, decomposing with respect to σ_{φ} yields

$$\mathcal{P}(\|\tilde{\eta}^{\varphi}(N^{\beta t})\|=k) \approx \int_{\alpha' t}^{\beta t} \mathcal{P}(\sigma_{\varphi}=dN^{s}, \ \|\tilde{\eta}^{\varphi}(N^{\beta t})\|=k). \tag{3.18}$$

The integrand can further be decomposed according to the countably many possible configurations $\chi \in \tilde{\Phi}$ at this first collision time N^{s}. By the strong Markov property of the process $\tilde{\eta}$ we can therefore continue with

$$= \int_{\alpha' t}^{\beta t} \sum_{\chi: \|\chi\|=n-1} \mathcal{P}(\sigma_{\varphi}=dN^{s}, \ \tilde{\eta}^{\varphi}(N^{s})=\chi) \ \mathcal{P}(\|\tilde{\eta}^{\chi}(N^{\beta t}-N^{s})\|=k). \tag{3.19}$$

Now we proceed as in step 4° of proof of Proposition 2.55: Besides an error term, χ can additionally be assumed to belong to the set $\tilde{\Phi}(n-1,s,2c)$ (defined in (3.12)). By 1°, in the expression (3.19) we may replace $\mathcal{P}(\|\tilde{\eta}^{\chi}(N^{\beta t}-N^{s})\|=k)$ by $\mathcal{P}(\|\tilde{\eta}^{\chi}(N^{\beta t})\|=k) + \varepsilon(t,\beta,\chi)$, uniformly for s in the interval $[\alpha'(t)t,\beta(t)t]$ and χ in $\tilde{\Phi}(n-1,s,2c)$. Moreover, by the induction hypothesis (replace *c* by *2c*)

$$\mathcal{P}(\|\tilde{\eta}^{\chi}(N^{\beta t})\|=k) = p_{n-1,k}(s/\beta t) + \varepsilon(t,\beta,\chi) \quad \text{as } t\to\infty,$$

uniformly for s in $[\alpha'(t)t,\beta(t)t]$ and χ in $\tilde{\Phi}(n-1,s,2c)$. Hence (3.19) can be continued with

(3.20) $$\approx \int_{\alpha' t}^{\beta t} \sum_{\chi\in\tilde\Phi} \mathcal{P}(\sigma_\varphi = dN^s,\ \tilde\eta^\varphi(N^s) = \chi \in \tilde\Phi(n-1,s,2c))\ p_{n-1,k}(s/\beta t).$$

Go back from $\tilde\Phi(n-1,s,2c)$ to *any* $(n-1)$-particle configurations, and use

$$p_{n-1,k}(s/\beta t) = \int_0^s p_{n-1,k}(dr/\beta t), \qquad 0\le s\le\beta.$$

Change the order of integration, and write (3.20) as

(3.21) $$\approx \int_0^{\beta t} p_{n-1,k}(dr/\beta t)\ \mathcal{P}\Big(\sigma_\varphi \in (N^{r\vee\alpha' t}, N^{\beta t})\Big).$$

By Proposition 2.55 and (3.10), we can continue with

$$\approx p_{n,n}(\alpha' t/(r\vee\alpha' t)) - p_{n,n}(\alpha' t/\beta t) = \int_{r\vee\alpha' t}^{\beta t} p_{n,n}(\alpha' t/ds),$$

uniformly in $r\in[\alpha'(t)t, \beta(t)t]$. Insert this in (3.21) and change again the order of integration to get $\approx \int_{\alpha' t}^{\beta\, t} p_{n,n}(\alpha' t/ds)\ p_{n-1,k}(s/\beta t)$. Again by Proposition 2.55, and by (3.10), this is equivalent to (3.17). This completes the induction step and hence finishes the proof of Proposition 3.13. ■

3.e) Multi-scale Spreading of Initial Points

Before we will generalize Proposition 3.13 (scaling limit) for collections of initial particles which spread apart at possibly different speed, we mention the following consequence of Lemma 2.26.

Lemma 3.22 (speed of spread). *Fix* $n\ge 1$ *and* $c>0$. *For* $t>0$, *let* $T(t)>0$ *and* $\varphi(t)\in\tilde\Phi$ *satisfy* $\|\varphi(t)\|=n$ *and*

(3.23) $$\|\zeta(1)-\zeta(2)\| \le [T(t)] + c\ell(T(t)) \quad \textit{if} \quad \varphi_{\zeta(1)}(t)\varphi_{\zeta(2)}(t)>0,\ \zeta(1)\ne\zeta(2).$$

Then $s(t)-2 > r(t) \ge -\infty$ *and* $s(t) \ge T(t) \underset{t\to\infty}{\longrightarrow} \infty$ *imply*

$$\tilde{\mathbf{P}}_{\varphi(t)}\Big(\tilde\eta(N^{s(t)}-N^{r(t)}) \in \tilde\Phi(j,s(t),2c) \text{ for some } j\ge 1\Big) \underset{t\to\infty}{\longrightarrow} 1.$$

Proof. Consider the opposite event. Then there is a pair of particles in $\tilde\eta(N^{s(t)}-N^{r(t)})$ with positions $\xi\ne\xi'$ satisfying $\xi-\xi' \notin \Xi(s(t),2c)$. Moreover, there must exist initial points $\zeta(1)$, $\zeta(2)$ with properties as in (3.23) and such that the arising walks do not coalesce by time $N^{s(t)}-N^{r(t)}$ and end up at ξ,ξ'. The probability of this event can be estimate from above by

$$\le \mathcal{P}\Big(Z(\zeta(1),N^{s(t)}-N^{r(t)}) - Z(\zeta(2),N^{s(t)}-N^{r(t)}) \notin \Xi(s(t),2c)\Big).$$

According to Lemma 2.26, for all t sufficiently large we may continue with $\le$ const $[s(t)]^{-2c}$ converging to 0 as $t\to\infty$. This finishes the proof since the number of pairs of initial particles remains bounded. ■

Remark 3.24. The fact that in $\tilde{\eta}$ particles coalesce instantaneously is not used in the previous proof. Hence, a statement analogously to Lemma 3.22 also holds for η, the coalescing random walk with delay. □

Assumption 3.25. Fix a constant $c\geq 1$, natural numbers $m\geq 1$, $n(1),\dots,n(m) \geq 1$, and let I be the index array $\{[i,u];\ 1\leq i\leq m,\ 1\leq u\leq n(i)\}$. For $t>0$, consider $0\leq\alpha(1,t)\leq\dots\leq\alpha(m,t)$, and t-dependent starting positions $\zeta(i,u) := \zeta(i,u,t)$, $[i,u]\in I$. For $i=1,\dots,m$ set $\varphi(i,t) := \delta^{\zeta(i,1)}+\dots+\delta^{\zeta(i,n(i))}$ and $\varphi(t) := \varphi(1,t)+\dots+\varphi(m,t)$. Recalling (2.25), on these starting systems we require $\varphi(t)\in\tilde{\Phi}$ and

(3.26) $$\zeta(i,u)-\zeta(j,v) \in \Xi(\alpha(j,t)t,c) \quad \text{whenever} \quad [i,u]\neq[j,v],\ i\leq j. \qquad \blacksquare$$

That is, $\alpha(j,t)$ describes the order of speed of spread within the j-th subsystem, and particles from different subsystems escape from each other with the order of speed of the particle with the bigger subsystem index.

Set $$p_{n;k}(\alpha;\beta) := 1\{n\geq k\}p_{n,k}(\alpha/\beta), \qquad n,k \geq 1,\quad 0\leq\alpha\leq\beta,$$
with the $p_{n,k}$ taken from (3.9) and (3.10). For $m>1$, $n(1),\dots,n(m),k \geq 1$, and $0\leq\alpha_1\leq\dots\leq\alpha_m\leq\beta$, recursively define

(3.27) $$p_{n(1),\dots,n(m);k}(\alpha_1,\dots,\alpha_m;\beta) := \sum_{l=1}^{\infty} p_{n(1),\dots,n(m-1);l}(\alpha_1,\dots,\alpha_{m-1};\alpha_m)\ p_{l+n(m);k}(\alpha_m;\beta).$$

Now we are prepared for

Proposition 3.28 **(multi-scale limit).** *In addition to Assumption* 3.25, *fix constants* $0<\beta_-\leq\beta_+<+\infty$. *For* $t>0$ *let* $\beta(t)\in[\beta_-,\beta_+]$. *Then*
$$\left|\tilde{\mathbf{P}}_{\varphi(t)}(\|\tilde{\eta}(N^{\beta(t)t})\|=k) - p_{n(1),\dots,n(m);k}(\alpha'(1,t),\dots,\alpha'(m,t);\beta(t))\right| \xrightarrow[t\to\infty]{} 0$$
where $1 \leq k \leq n(1)+\dots+n(m)$ *and* $\alpha' := \alpha\wedge\beta$.

Proof. In the case $m=1$ the claim is our *scaling limit Proposition* 3.13. We shall prove the result by induction over m. Suppose the claim holds for some $m-1 \geq 1$. We may assume that $\alpha(m,t)$ is bounded away from 0. Introduce the event
$$E(t) := \left\{\sigma_{[i,u],[m,v]} > N^{\alpha'(m,t)t},\quad [i,u]\neq[m,v],\ i\leq m\right\},$$
i.e. that there is not a single collision by time $N^{\alpha'(m,t)t}$ with a particle

of the m-th subsystem. First observe that for the complement $\mathcal{C}E(t)$ of $E(t)$

$$\mathcal{P}(\mathcal{C}E(t)) \leq \sum_{[i,u]\neq[m,v]} 1\{i\leq m\}\ \mathcal{P}\left\{\sigma_{[i,u],[m,v]} \leq N^{\alpha'(m,t)t}\right\}$$

holds, where each of these finitely summands converges to 0 as $t\to\infty$ by our Proposition 2.55. Therefore it is enough to study

$$\sum_{\psi,\chi:\ \chi\wedge\psi=0,\ \|\chi\|=n(m)} \mathcal{P}\left(E\cap\left\{\tilde{\eta}^{\varphi-\varphi(m)}(N^{\alpha'(m)t})=\psi,\ \tilde{\eta}^{\varphi(m)}(N^{\alpha'(m)t})=\chi,\ \|\tilde{\eta}^{\varphi}(N^{\beta t})\|=k\right\}\right)$$

(for convenience, in notation we suppressed the t-dependence at several places), where we could restrict to $\chi,\psi\in\tilde{\Phi}$ "disjoint" since $\tilde{\eta}^{\varphi-\varphi(m)}$ and $\tilde{\eta}^{\varphi(m)}$ do not interact before time $N^{\alpha'(m)t}$ under $E(t)$, and hence on this event

$$\tilde{\eta}^{\varphi}(N^{\alpha'(m)t}) = \tilde{\eta}^{\varphi-\varphi(m)}(N^{\alpha'(m)t}) + \tilde{\eta}^{\varphi(m)}(N^{\alpha'(m)t}).$$

We can use the Markov property to write the summands in the sum above as

$$\mathcal{P}\left(E\cap\left\{\tilde{\eta}^{\varphi-\varphi(m)}(N^{\alpha'(m)t})=\psi,\ \tilde{\eta}^{\varphi(m)}(N^{\alpha'(m)t})=\chi\right\}\right)\ \mathcal{P}\left(\|\tilde{\eta}^{\psi+\chi}(N^{\beta t}-N^{\alpha'(m)t})\|=k\right).$$

Set $\|\psi\|=:l$. Applying Proposition 3.13, we can conclude that for $\omega = \omega(t) \in \tilde{\Phi}(l+n(m),\alpha'(m,t)t,2c)$

$$\lim_{t\to\infty} \left|\mathcal{P}\left(\|\tilde{\eta}^{\omega}(N^{\beta t}-N^{\alpha'(m)t})\|=k\right) - p_{l+n(m);k}(\alpha'(m);\beta)\right| = 0.$$

Next we use that the $\omega=\psi+\chi$ excluded so far do not contribute, by Lemma 3.22 (with $r\equiv-\infty$). Hence, we need to find the limit as $t\to\infty$ of

$$\sum_l \mathcal{P}\left(\|\tilde{\eta}^{\varphi-\varphi(m)}(N^{\alpha'(m)t})\|=l\right)\ p_{l+n(m);k}(\alpha'(m);\beta).$$

Since l takes on only finitely many values, we can use the induction hypothesis and the definition (3.27) to complete the proof. ∎

4. The Time Picture of Components

The purpose of this section is to prove the Theorems 4.a) and b) which we shell need later in the proof of the Theorems 1-3. The main tools will be comparison arguments and moment estimates. We start by developing these tools in 4.a-c), the final proofs of the Theorems 4.a),b) are in 4.d),e), respectively.

4.a) Comparison and Coupling Techniques

Because in the claims of our theorems the limiting expressions do not depend on the diffusion coefficient $g\in\mathcal{G}^0$ (defined in (1.2)), the basic idea to get *universality* in g is to *compare* the general interacting system corresponding to a given $g\in\mathcal{G}^0$ with adequate interacting Fisher-Wright systems. In fact, since such g is (strictly) positive on $(0,1)$, and g is Lipschitz continuous, for each $\varepsilon\in(0,1/2)$ there exist positive constants b^ε and b such that

(4.1) $$g^\varepsilon := b^\varepsilon \ell^\varepsilon \le g \le b\ell \quad \text{with} \quad \ell^\varepsilon(r) := (r-\varepsilon)^+(1-\varepsilon-r)^+, \quad 0\le r\le 1.$$

Of course, the majorizing term $b\ell$ determines a Fisher-Wright system on the interval $[0,1]$ (recall (1.3)).

On the other hand, note that $g^\varepsilon = b^\varepsilon \ell^\varepsilon$ belongs to $\mathcal{G}\supset\mathcal{G}^0$, see (2.57). If however for the moment we restrict our consideration to those initial laws $\mu^\varepsilon\in\mathfrak{P}$ satisfying

(4.2) $$\mu^\varepsilon([\varepsilon,1-\varepsilon]^\Xi)=1,$$

then we can identify the interacting diffusion system $\mathfrak{X}=[\mathfrak{X},\mathbb{P}^{g^\varepsilon}_{\mu^\varepsilon},\mu^\varepsilon\in\mathfrak{P}]$ on $[0,1]$ with an *interacting Fisher-Wright system* $\mathfrak{X}^\varepsilon=[\mathfrak{X}^\varepsilon,\mathbb{P}^{g^\varepsilon}_\mu,\mu\in\mathfrak{P}^\varepsilon]$ *on* $[\varepsilon,1-\varepsilon]$, where $\mathfrak{P}^\varepsilon$ denotes the set of all probability laws on $[\varepsilon,1-\varepsilon]^\Xi$). In fact, the linear transformation $\mathbf{L}^\varepsilon$: $[\varepsilon,1-\varepsilon]^\Xi \mapsto [0,1]^\Xi$ defined by

(4.3) $$(\mathbf{L}^\varepsilon \mathbf{x})_\xi := (\mathbf{x}_\xi-\varepsilon)/(1-2\varepsilon), \qquad \mathbf{x}\in[\varepsilon,1-\varepsilon]^\Xi,\ \xi\in\Xi$$

maps this $\mathfrak{X}^\varepsilon$ into an interacting Fisher-Wright system $\mathfrak{Y}=[\mathfrak{Y},\mathbb{P}^{b^\varepsilon}_\mu,\mu\in\mathfrak{P}]$ on $[0,1]$ with diffusion constant b^ε, since

(4.4) $$g^\varepsilon(\mathbf{x}_\xi) = (1-2\varepsilon)^2 b^\varepsilon \ell((\mathbf{L}^\varepsilon\mathbf{x})_\xi), \qquad \mathbf{x}\in[\varepsilon,1-\varepsilon]^\Xi,\ \xi\in\Xi.$$

In the case of an interacting diffusion system with diffusion coefficient $g^\varepsilon = b^\varepsilon\ell^\varepsilon$, we need therefore a method to reduce proofs to initial laws $\mu^\varepsilon\in\mathfrak{P}$ with (4.2). We use the following *concept of coupling* of interacting diffusions (in Subsection 5.b) we shall deepen this concept).

Definition 4.5 **(coupling principle).** Fix $g\in\mathcal{G}$ and two (possibly different) initial laws $\mu,\nu\in\mathfrak{P}$. Let Γ be a distribution on $[0,1]^\Xi\times[0,1]^\Xi$ with marginals μ,ν. Choose $[\mathfrak{X}(0),\mathfrak{Y}(0)]$ according to Γ, and solve (1.7) separately for μ and ν, but using the *same* collection $\mathbf{w}=\{w_\xi\}_{\xi\in\Xi}$ of the driving Wiener processes (recall that equation (1.7) has a unique *strong* solution). Then the bivariate process $[\mathfrak{X},\mathfrak{Y}]$ is called the *coupling of the interacting diffusion systems* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *with joint initial law* Γ. ▮

We will use this coupling principle to show the following

Lemma 4.6. *Fix $\mu\in\mathfrak{P}$ and $\varepsilon\in(0,1/2)$. Then there is a $\mu^{\varepsilon}\in\mathfrak{P}$ satisfying the condition (4.2), and a law Γ on $[0,1]^{\Xi}\times[0,1]^{\Xi}$ with marginals μ,μ^{ε} and such that for the coupling $[\mathfrak{X},\mathfrak{Y}]$ with joint initial law Γ,*

$$\mathbb{E}^{q}_{\Gamma}|x_{\xi}(t) - \eta_{\xi}(t)| \le \varepsilon, \quad t\ge 0,\ \xi\in\Xi,\ q\in\mathcal{G}. \tag{4.7}$$

Proof. Let be given μ,ε as in the lemma. Realize $\mathfrak{X}(0)\in[0,1]^{\Xi}$ according to the law μ. Define $\mathfrak{Y}(0)$ by projecting components $x_{\xi}(0)$ of $\mathfrak{X}(0)$ with values in $[0,\varepsilon]$ or $[1-\varepsilon,1]$ to $\eta_{\xi}(0)=\varepsilon$ or $\eta_{\xi}(0)=1-\varepsilon$, respectively, and set $\eta_{\xi}(0)=x_{\xi}(0)$ otherwise, $\xi\in\Xi$. Let Γ and μ^{ε} denote the laws of $[\mathfrak{X}(0),\mathfrak{Y}(0)]$ and $\mathfrak{Y}(0)$, respectively. Then

$$\mathbb{E}_{\Gamma}|x_{\xi}(0)-\eta_{\xi}(0)| \le \varepsilon \quad \text{for all } \xi\in\Xi. \tag{4.8}$$

Pass to the coupling $[\mathfrak{X},\mathfrak{Y}]$ with joint initial law Γ according to the Definition 4.5. Applying the expectation formula (2.59), we get

$$\mathbb{E}^{q}_{\Gamma}|x_{\xi}(t) - \eta_{\xi}(t)| \le \textstyle\sum_{\zeta} p(t,\xi,\zeta)\, \mathbb{E}_{\Gamma}|x_{\zeta}(0) - \eta_{\zeta}(0)| \le \varepsilon, \quad t\ge 0,\ \xi\in\Xi,\ q\in\mathcal{G}. \qquad \blacksquare$$

Remark 4.9. If μ belongs to $\mathfrak{I}_{\theta}$, then μ^{ε} constructed in the previous proof is also shift-invariant and ergodic, and its density θ^{ε} tends to θ as $\varepsilon\to 0$. $\square$

4.b) Comparison of Moments

Now we come to the following basic comparison formulas for general systems by Fisher-Wright systems (recall (4.1)).

Proposition 4.10 (comparison of moments). *Fix $q\in\mathcal{G}$, $t\ge 0$ and $\varphi\in\Phi$. Then the following inequalities hold.*

(j) *If $b>0$ satisfies $q\le b\ell$ then:* $\mathbb{E}^{q}_{\mathbf{z}}(\mathfrak{X}(t))^{\varphi} \le \mathbb{E}^{b}_{\mathbf{z}}(\mathfrak{X}(t))^{\varphi}$, $\mathbf{z}\in[0,1]^{\Xi}$.

(jj) *If $\varepsilon\in[0,1/2)$ and $b^{\varepsilon}>0$ satisfy $q^{\varepsilon}:=b^{\varepsilon}\ell^{\varepsilon}\le q$ (with $\ell^{0}:=\ell$) then:*

$$0 \le \mathbb{E}^{q^{\varepsilon}}_{\mathbf{z}}(\mathfrak{X}(t))^{\varphi} \le \mathbb{E}^{q}_{\mathbf{z}}(\mathfrak{X}(t))^{\varphi}, \quad \mathbf{z}\in[\varepsilon,1-\varepsilon]^{\Xi}.$$

Proof. We will give only the proof of the statement (jj) since (j) can be considered as (jj) reversed in the special case $\varepsilon=0$.

Recall the semigroup U^{q} with generator $\mathfrak{G}^{q}$ related to $\mathfrak{X}$, which was defined in (2.60). Take $q,t,\varepsilon,b^{\varepsilon}$, as assumed in (jj). We have to show that

$$U^{q^{\varepsilon}}_{t} h_{\varphi}(\mathbf{z}) \le U^{q}_{t} h_{\varphi}(\mathbf{z}), \quad \mathbf{z}\in[\varepsilon,1-\varepsilon]^{\Xi}, \tag{4.11}$$

where again $h_\varphi(\mathbf{z})=\mathbf{z}^\varphi$.

As pointed out in the previous subsection, the interacting diffusion $\mathfrak{X}$ with diffusion coefficient $g^\varepsilon=\mathfrak{b}^\varepsilon \ell^\varepsilon$ and initial state in $[\varepsilon,1-\varepsilon]^\Xi$ can be considered as an interacting Fisher-Wright diffusion $\mathfrak{X}^\varepsilon$ on $[\varepsilon,1-\varepsilon]$. Moreover, the linear transformation $\mathbf{L}^\varepsilon$ defined in (4.3) transforms $\mathfrak{X}^\varepsilon$ to an interacting Fisher-Wright system $\mathfrak{Y}$ on $[0,1]$ with diffusion coefficient $\mathfrak{b}^\varepsilon$. Hence for $s\geq 0$,

$$(4.12)\qquad U_s^{g^\varepsilon} h_\varphi(\mathbf{z}) = \mathbb{E}_{\mathbf{z}}^{g^\varepsilon}(\mathfrak{X}(s))^\varphi = \mathbb{E}_{\mathbf{L}^\varepsilon \mathbf{z}}^{\mathfrak{b}^\varepsilon}(\mathbf{H}^\varepsilon \mathfrak{Y}(s))^\varphi,$$

where $\mathbf{H}^\varepsilon$: $[0,1]^\Xi \mapsto [\varepsilon,1-\varepsilon]^\Xi$ is the inverse mapping to $\mathbf{L}^\varepsilon$, that is

$$(4.13)\qquad (\mathbf{H}^\varepsilon \mathbf{x})_\xi = (1-2\varepsilon)(\mathbf{x}_\xi+\varepsilon), \qquad \mathbf{x}\in[0,1]^\Xi,\ \xi\in\Xi.$$

We want to show:

$$(4.14)\qquad \mathbf{z}\mapsto U_s^{g^\varepsilon} h_\varphi(\mathbf{z}),\ \mathbf{z}\in[\varepsilon,1-\varepsilon]^\Xi, \text{ is } convex \text{ in each component } \mathbf{z}_\xi,\ \xi\in\Xi.$$

At the r.h.s. of (4.12), extract $(1-2\varepsilon)^{\|\varphi\|}$ and compute the binomial in each factor $(\eta_\xi(s)+\varepsilon)^{\varphi_\xi}$ to recognize that the r.h.s. can be written as a *non-negative* combination of terms of the form

$$(4.15)\qquad \mathbb{E}_{\mathbf{L}^\varepsilon \mathbf{z}}^{\mathfrak{b}^\varepsilon}(\mathfrak{Y}(s))^\psi = U_s^{\mathfrak{b}^\varepsilon} h_\psi(\mathbf{L}^\varepsilon \mathbf{z}) \quad \text{with} \quad \psi\in\Phi.$$

For $\mathfrak{b}>0$, by the *duality* Lemma 3.2:

$$U_s^{\mathfrak{b}} h_\psi(\mathbf{y}) = \mathbb{E}_{\mathbf{y}}^{\mathfrak{b}}(\mathfrak{X}(s))^\psi = \mathbf{E}_\psi \mathbf{y}^{\eta(s)}, \qquad \mathbf{y}\in[0,1]^\Xi,\ \psi\in\Phi,\ s\geq 0.$$

Therefore,

$$(\partial/\partial \mathbf{y}_\xi)^2 U_s^{\mathfrak{b}} h_\psi(\mathbf{y}) = (\partial/\partial \mathbf{y}_\xi)^2 \mathbf{E}_\psi \mathbf{y}^{\eta(s)} = \mathbf{E}_\psi \eta_\xi(s)(\eta_\xi(s)-1)\mathbf{y}^{\eta(s)-2\delta^\xi} \geq 0, \qquad \xi\in\Xi.$$

Hence, (for fixed $\mathfrak{b},s,\psi$) the function $\mathbf{y}\mapsto U_s^{\mathfrak{b}} h_\psi(\mathbf{y})$, $\mathbf{y}\in[0,1]^\Xi$, is convex in each component $\mathbf{y}_\xi$, $\xi\in\Xi$. Consequently, the r.h.s of (4.15) is convex in each component $\mathbf{z}_\xi$, $\xi\in\Xi$, hence, as a non-negative combination, the r.h.s. of (4.12) is as well. This proves (4.14).

Thus, $(\mathfrak{G}^{g^\varepsilon}-\mathfrak{G}^{g})U_s^{g^\varepsilon} h_\varphi \leq 0$ on $[\varepsilon,1-\varepsilon]^\Xi$ since $g^\varepsilon - g \leq 0$. Then the *integration by parts formula*

$$(4.16)\qquad U_t^{g^\varepsilon} - U_t^{g} = \int_0^t ds\ U_{t-s}^{g}(\mathfrak{G}^{g^\varepsilon}-\mathfrak{G}^{g})U_s^{g^\varepsilon}$$

(see e.g.Liggett (1985), p. 367) yields (4.11). This finishes the proof. ■

Using part of the argument of the previous proof we get

Lemma 4.17. *Fix* $g\in\mathcal{G}$, $t\geq 0$, $\zeta,\zeta'\in\Xi$. *If* $\varepsilon\in[0,1/2)$, $\mathfrak{b}^\varepsilon>0$, *and* $g^\varepsilon=\mathfrak{b}^\varepsilon \ell^\varepsilon \leq g$, *then*

$$(4.18)\qquad 0 \le \mathbb{E}_{\mathbf{z}}^{q}\mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(t),\mathfrak{x}_{\zeta'}(t)) \le \mathbb{E}_{\mathbf{z}}^{q^{\varepsilon}}\mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(t),\mathfrak{x}_{\zeta'}(t)), \qquad \mathbf{z}\in[\varepsilon,1-\varepsilon]^{\Xi},$$

where

$$(4.19)\qquad \mathcal{F}^{\varepsilon}(r,r') := 2^{-1}((r-\varepsilon)^{+}(1-\varepsilon-r')^{+} + (r'-\varepsilon)^{+}(1-\varepsilon-r)^{+}), \quad 0\le r,r'\le 1.$$

Proof. First note that $\mathbf{x} \mapsto h(\mathbf{x}) := \mathcal{F}^{\varepsilon}(\mathbf{x}_{\zeta},\mathbf{x}_{\zeta'})$, $\mathbf{x}\in[0,1]^{\Xi}$, belongs to $\mathbf{C}[[0,1]^{\Xi}]$, and we may apply the U^{q}-semigroup to it. Then

$$(4.20)\qquad U_{s}^{q^{\varepsilon}}h(\mathbf{z}) = \mathbb{E}_{\mathbf{z}}^{q^{\varepsilon}}\mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta'}(s)), \qquad s\ge 0,$$

which is a *concave* function in each component $\mathbf{z}_{\xi}\in[\varepsilon,1-\varepsilon]$, $\xi\in\Xi$. In fact, $\mathfrak{X}(s)$ belongs to $[\varepsilon,1-\varepsilon]^{\Xi}$ with $\mathbb{P}_{\mathbf{z}}^{q^{\varepsilon}}$-probability one (for s fixed). Moreover, by the symmetric definition (4.19) of $\mathcal{F}^{\varepsilon}$,

$$(4.21)\qquad \mathcal{F}^{\varepsilon}(r,r') = \mathcal{F}^{0}(r,r') - \ell(\varepsilon), \qquad r,r' \in [\varepsilon,1-\varepsilon].$$

Hence, the $\mathcal{F}^{\varepsilon}$-expression at the r.h.s. of (4.20) can easily be computed leading to a constant and first moment expressions except the term $-\mathbb{E}_{\mathbf{z}}^{q^{\varepsilon}}\mathfrak{x}_{\zeta}(s)\mathfrak{x}_{\zeta'}(s)$. But first moments are linear in $\mathbf{z}$ (see (2.59)) whereas the latter summand is concave in each component $\mathbf{z}_{\xi}$ according to (4.14). Hence, $(\mathfrak{G}^{q^{\varepsilon}}-\mathfrak{G}^{q})U_{s}^{q^{\varepsilon}}h \ge 0$ on $[\varepsilon,1-\varepsilon]^{\Xi}$, and again the integration by parts formula (4.16) yields the claim. ∎

4.c) Some Uniform Moment Estimates

A crucial ingredient for the proof of Theorem 4 are moment estimates which we will derive in this subsection (see Proposition 4.24 below).

We start with some preparations. From (2.59) and Lemma 2.16 we immediately obtain the following *first moment estimate.*

Lemma 4.22. *Fix $\zeta,\zeta'\in\Xi$. There exist constants $c,C>0$ such that*

$$(4.23)\qquad \left|\mathbb{E}_{\mathbf{z}}^{q}\mathfrak{x}_{\zeta}(t) - \mathbb{E}_{\mathbf{z}}^{q}\mathfrak{x}_{\zeta'}(t)\right| \le C\, e^{-ct}, \qquad \mathbf{z}\in[0,1]^{\Xi},\ t\ge 0,\ q\in\mathcal{G}.$$

We will need the following (partial) generalization of Lemma 2.61. Recall the notations in (4.1) and (4.19).

Proposition 4.24 **(a speed of convergence estimate).** *Fix $\zeta,\zeta'\in\Xi$ and $q\in\mathcal{G}$. Then there are constants $c,C>0$ such that for $\varepsilon\in[0,1/2)$ and $b^{\varepsilon}>0$ with $q^{\varepsilon}=b^{\varepsilon}\ell^{\varepsilon}\le q$,*

$$(4.25)\qquad \mathbb{E}_{\mathbf{z}}^{q}\,\mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(t),\mathfrak{x}_{\zeta'}(t)) \le C(e^{-ct} + 1/b^{\varepsilon}\log t), \qquad \mathbf{z}\in[\varepsilon,1-\varepsilon]^{\Xi},\ t>1.$$

Proof. By Lemma 4.17, we loose no generality in supposing that $q^{\varepsilon}=b^{\varepsilon}\ell^{\varepsilon}$. First we shall deal with the particular case $\zeta=\zeta'$. We even want to show

$$\mathbb{E}_{\mathbf{z}}^{g^{\varepsilon}} \ell^{\varepsilon}(\mathfrak{x}_{\xi}(t)) \leq \text{const}/b^{\varepsilon}\log t, \quad \mathbf{z}\in[\varepsilon,1-\varepsilon]^{\Xi},\ t>1,\ \xi\in\Xi. \tag{4.26}$$

By (4.4) and the transformation formula (4.12),

$$\mathbb{E}_{\mathbf{z}}^{g^{\varepsilon}} \ell^{\varepsilon}(\mathfrak{x}_{\xi}(t)) = (1-2\varepsilon)^2\, \mathbb{E}_{\mathbf{y}}^{b^{\varepsilon}} \ell(\mathfrak{x}_{\xi}(t)) \leq \mathbb{E}_{\mathbf{y}}^{b^{\varepsilon}} \ell(\mathfrak{x}_{\xi}(t)).$$

with $\mathbf{y}:=\mathbf{L}^{\varepsilon}\mathbf{z}$. By the *duality* Lemma 3.2 (with coalescing rate b^{ε})

$$\mathbb{E}_{\mathbf{y}}^{b^{\varepsilon}} \ell(\mathfrak{x}_{\xi}(t)) = \mathbf{E}_{\delta_{\xi}} \mathbf{y}^{\eta(t)} - \mathbf{E}_{2\delta_{\xi}} \mathbf{y}^{\eta(t)}. \tag{4.27}$$

Split the last expectation according to $\|\eta(t)\|=1$ or 2. Then in the second case the restricted expectation can be bounded from above by the probability $q_{2,2}(t)$ of not coalescing by time t, which is independent of $\xi,\mathbf{y}$, and of order $1/b^{\varepsilon}\log t$ according to Lemma 3.3. Using the coupling Convention 3.5, the remaining expressions can be written as

$$\mathcal{E}\mathbf{y}_{Z(\xi,t)}\left(1 - \mathcal{P}\left\{Z(\xi) \text{ and } Z'(\xi) \text{ coalesce by time } t \,\middle|\, Z(\xi)\right\}\right)$$

$$= \mathcal{E}\mathbf{y}_{Z(\xi,t)}\, \mathcal{P}\left\{Z(\xi) \text{ and } Z'(\xi) \text{ do not coalesce by time } t \,\middle|\, Z(\xi)\right\},$$

where $Z(\xi)$ and $Z'(\xi)$ are independent random walks. Thus, this can also be estimated from above by $\leq q_{2,2}(t)$. Summarizing, (4.26) holds, and the lemma is true in the particular case $\zeta=\zeta'$.

In order to remove the restriction to $\zeta=\zeta'$ by comparison, we will show

$$\left|\mathbb{E}_{\mathbf{z}}^{g^{\varepsilon}} \mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(t),\mathfrak{x}_{\zeta'}(t)) - \mathbb{E}_{\mathbf{z}}^{g^{\varepsilon}} \ell^{\varepsilon}(\mathfrak{x}_{\zeta}(t))\right| \leq C(e^{-ct} + 1/b^{\varepsilon}\log t), \tag{4.28}$$

$\mathbf{z}\in[\varepsilon,1-\varepsilon]^{\Xi}$, $t>1$ (for fixed ζ,ζ' and constants c,C independent of $\varepsilon,b^{\varepsilon}$). By the formula (4.21), the l.h.s. of (4.28) equals $\left|\mathbb{E}_{\mathbf{z}}^{g^{\varepsilon}}[\mathfrak{x}_{\zeta}^2(t) - \mathfrak{x}_{\zeta}(t)\mathfrak{x}_{\zeta'}(t)]\right|$. Using twice Lemma 2.58 in the case n=2, this can be estimated from above by

$$\leq \left|a_{\mathbf{z},\zeta}(t)-a_{\mathbf{z},\zeta'}(t)\right| + \int_0^t ds \sum_{\xi} p(t-s,\zeta,\xi)\left|p(t-s,\zeta,\xi)-p(t-s,\zeta',\xi)\right| \mathbb{E}_{\mathbf{z}}^{g^{\varepsilon}} g^{\varepsilon}(\mathfrak{x}_{\xi}(s))$$

By the Lemmas 4.22 and 2.16, we can continue with

$$\leq Ce^{-ct} + \int_0^t ds\ Ce^{-c(t-s)} \sup_{\xi\in\Xi}\mathbb{E}_{\mathbf{z}}^{g^{\varepsilon}} g^{\varepsilon}(\mathfrak{x}_{\xi}(s)).$$

But $g^{\varepsilon}=b^{\varepsilon}\ell^{\varepsilon}$ where $b^{\varepsilon} \leq 4\sup_{0<r<1} g(r) = \text{const}$. Thus it suffices to show that

$$\int_0^t ds\ e^{-c(t-s)} \sup_{\xi\in\Xi}\mathbb{E}_{\mathbf{z}}^{g^{\varepsilon}} \ell^{\varepsilon}(\mathfrak{x}_{\xi}(s)) \leq \text{const}/b^{\varepsilon}\log t. \tag{4.29}$$

For $s>2$, according to (4.26) we can bound the supremum by $\leq \text{const}/b^{\varepsilon}\log s$. Then continue by splitting the integral $\int_2^t ds\ e^{-c(t-s)}/\log s$ according to

$$\int_{t/2}^t ds\ e^{-c(t-s)}/\log s \leq (1/\log(t/2)) \int_0^{\infty} ds\ e^{-cs} \leq \text{const}/\log t,$$

$$(4.30) \qquad \int_2^{t/2} ds\; e^{-c(t-s)}/\log s \le e^{-ct/2} \int_2^{t/2} ds\; /\log s.$$

But, for all $m,n \ge 0$,

$$(4.31) \qquad \int_1^t ds\; s^n/\log^m s \;\sim\; (n+1)^{-1}\, t^{n+1}/\log^m t \quad \text{as} \quad t\to\infty.$$

Therefore we may continue to estimate (4.30) by

$$\le \text{const}\; e^{-ct/2}\, t/\log t \le \text{const}/\log t, \qquad t>2.$$

If $s\le 2$, estimate the supremum in (4.29) by 1 and use $e^{-c(t-s)} \le e^{2c}e^{-ct} \le$ $\text{const}/\log t$, $t>1$. Combining these estimates gives (4.28), and the proof is finished. ∎

Based on the previous result we will get some *fourth moment estimate* (recall the notations in (4.1) and (4.19)):

Lemma 4.32. *Fix $\zeta,\zeta'\in\Xi$ and $g\in\mathcal{G}$. Then there exists a constant $C>0$ such that for all $\varepsilon\in[0,1/2)$ and $b^\varepsilon>0$ satisfying $g^\varepsilon = b^\varepsilon \ell^\varepsilon \le g$,*

$$(4.33) \qquad \mathbb{E}_{\mathbf{z}}^{g}\left(\int_0^t ds\; \mathcal{F}^\varepsilon(\mathfrak{x}_\zeta(s),\mathfrak{x}_{\zeta'}(s))\right)^2 \le C(1 + t/b^\varepsilon \log t)^2, \qquad \mathbf{z}\in[\varepsilon,1-\varepsilon]^\Xi,\; t>2.$$

Proof. We start with a technical point. Since $\mathcal{F}^\varepsilon$ (introduced in (4.19)) is uniformly bounded, we may obviously restrict our consideration to

$$\mathbb{E}_{\mathbf{z}}^{g} \int_1^{t-1} ds\; \mathcal{F}^\varepsilon(\mathfrak{x}_\zeta(s),\mathfrak{x}_{\zeta'}(s)) \int_{s+1}^t dr\; \mathcal{F}^\varepsilon(\mathfrak{x}_\zeta(r),\mathfrak{x}_{\zeta'}(r)).$$

Using the Markov property of $\mathfrak{X}$, we can rewrite this as follows:

$$= \int_1^{t-1} ds\; \mathbb{E}_{\mathbf{z}}^{g}\, \mathcal{F}^\varepsilon(\mathfrak{x}_\zeta(s),\mathfrak{x}_{\zeta'}(s)) \int_{s+1}^t dr\; \mathbb{E}_{\mathfrak{X}(s)}^{g} \mathcal{F}^\varepsilon(\mathfrak{x}_\zeta(r-s),\mathfrak{x}_{\zeta'}(r-s)).$$

By Lemma 4.17 we may replace g by g^ε. Then $\mathfrak{X}(s)\in[\varepsilon,1-\varepsilon]$, $\mathbb{P}_{\mathbf{z}}^{g^\varepsilon}$-a.s. Thus, with the help of Proposition 4.24, we may bound the expectations to continue with

$$(4.34) \qquad \le \text{const} \int_1^{t-1} ds\; (e^{-cs} + 1/b^\varepsilon \log s) \int_1^{t-s} dr\; (e^{-cr} + 1/b^\varepsilon \log r).$$

Hence, again by symmetry, it suffices to estimate the quadratic expression

$$\left(\int_1^{t-1} ds\; (e^{-cs} + 1/b^\varepsilon \log s)\right)^2$$

which is by (4.31) of the desired order. ∎

4.d) Clumps Sticking at the Boundary

Before we will come to the proof of Theorem 4.a), we prepare an important tool and study the occupation time of a certain function of components (recall (4.1) and (4.19)):

Lemma 4.35. *Fix* ζ,ζ', $\mu\in\mathfrak{P}$, $g\in\mathcal{G}$. *Let* $\mu([\varepsilon,1-\varepsilon]^{\Xi})=1$ *and* $g^{\varepsilon}=\mathfrak{b}^{\varepsilon}\ell^{\varepsilon}\le g$ *for some* $\mathfrak{b}^{\varepsilon}>0$ *and* $\varepsilon\in[0,1/2)$. *Then*

$$t^{-1}\int_0^t ds\ \mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta'}(s)) \underset{t\to\infty}{\longrightarrow} 0 \quad \mathbb{P}_{\mu}^{g}\text{-a.s.} \tag{4.36}$$

Proof. By Lemma 4.32 we have

$$\mathbb{E}_{\mu}^{g}\left(t^{-1}\int_0^t ds\ \mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta'}(s))\right)^2 \le C(1/t+1/\mathfrak{b}^{\varepsilon}\log t)^2 \le \text{const}/(\log t)^2 \tag{4.37}$$

as $t\to\infty$ (since ε is fixed). Then by standard arguments, see e.g. the proof of Theorem 2 in Cox and Griffeath (1983), the statement follows. In fact, for fixed $r>1$ and $\delta>0$, from (4.37) we conclude

$$\mathbb{P}_{\mu}^{g}\left(r^{-n}\int_0^{r^n} ds\ \mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta'}(s)) > \delta\right) \le \text{const } n^{-2}, \quad n\ge n_r, \tag{4.38}$$

for some n_r. Then by Borel-Cantelli, (4.36) holds along the sequence $t=r^n$. By monotonicity in t of the occupation time, for $r^n\le t\le r^{n+1}$ we have

$$\begin{aligned} r^{-1}\left(r^{-n}\int_0^{r^n} ds\ \mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta'}(s))\right) &\le t^{-1}\int_0^t ds\ \mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta'}(s)) \\ &\le r\left(r^{-n-1}\int_0^{r^{n+1}} ds\ \mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta'}(s))\right), \end{aligned}$$

and letting first $t\to\infty$ and then $r\to 1$ (through a countable sequence) the statement (4.36) follows. ■

Now we are in a position to complete the

Proof of Theorem 4.a). We shell first prove the stronger a.s. statement under the additional Condition 1.20. Later we point out how to drop the assumption to get the weaker statement of the theorem. Fix $\zeta,\zeta'\in\Xi$, $\mu\in\mathfrak{P}$, $g\in\mathcal{G}^0$, and $0<\delta<1/2$.

First we work with the additional condition (1.22), i.e. we assume that $\mu([\varepsilon_0,1-\varepsilon_0]^{\Xi})=1$ for some $0<\varepsilon_0<1/2$. As already discussed in the beginning of Subsection 4.a), to each $\varepsilon\in(0,\varepsilon_0\wedge\delta]$ we may choose a constant $\mathfrak{b}^{\varepsilon}>0$ such that $g^{\varepsilon}=\mathfrak{b}^{\varepsilon}\ell^{\varepsilon}\le g$. Assume for the moment that $\zeta=\zeta'$, and consider the contribution to the integral in (4.36) on the set of s-values such that $\{\delta\le\mathfrak{x}_{\zeta}(s)\le 1-\delta\}$. On this set, $\mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta}(s))=\ell^{\varepsilon}(\mathfrak{x}_{\zeta}(s))$ is bounded away from 0, and we get the a.s. convergence (1.19) for $m=1$. Hence,

$$t^{-1}\int_0^t ds\ 1\{\mathfrak{x}_{\zeta}(s)<\delta \text{ or } \mathfrak{x}_{\zeta}(s)>1-\delta\}\ 1\{\mathfrak{x}_{\zeta'}(s)<\delta \text{ or } \mathfrak{x}_{\zeta'}(s)>1-\delta\} \underset{t\to\infty}{\longrightarrow} 1, \tag{4.39}$$

$\zeta,\zeta'\in\Xi$, $\mathbb{P}_{\mu}^{g}$-a.s. Now we repeat the arguments with ζ,ζ'. Namely: $\mathcal{F}^{\varepsilon}(\mathfrak{x}_{\zeta}(s),\mathfrak{x}_{\zeta'}(s))$

is bounded away from 0 on the set $\{\delta\leq\mathfrak{x}_\zeta(s),\ \mathfrak{x}_{\zeta'}(s)\leq 1-\delta\}$, and we conclude by symmetry that

$$t^{-1}\int_0^t ds\ 1\{\mathfrak{x}_\zeta(s)<\delta' \text{ or } \mathfrak{x}_{\zeta'}(s)>1-\delta'\}\ 1\{\mathfrak{x}_{\zeta'}(s)<\delta' \text{ or } \mathfrak{x}_\zeta(s)>1-\delta'\} \underset{t\to\infty}{\longrightarrow} 1 \quad \mathbb{P}_\mu^{g}\text{-a.s.}$$

Combined with (4.39), we get (1.19) for $m=2$. The cases $m>2$ follow by exclusion/inclusion from the cases $m=1,2$. Summarizing, we showed the $\mathbb{P}_\mu^{g}$-a.s. convergence (1.19) under the additional assumption (1.22).

Now we come to the proof under the condition (1.21). Here we can find a constant $b^0>0$ such that $b^0\ell\leq g$. Then we proceed in the same manner as above but with $\varepsilon=\varepsilon_0=0$. Summarizing, (1.19) holds $\mathbb{P}_\mu^{g}$-a.s. hence in $\mathbb{P}_\mu^{g}$-probability, provided that Condition 1.20 is fulfilled.

Finally, for general μ and g (i.e. without either (1.21) or (1.22) holding), we obtain the stochastic convergence by using the coupling Lemma 4.6 to reduce the assertion to the case where (1.22) is assumed. This finishes the proof of Theorem 4.a). ■

4.e) Law of Large Numbers and Oscillations: Proof of Theorem 4.b)

For the proof of (1.24) we need the following basic fact concerning the *convergence of expectations.*

Lemma 4.40. *Fix $g\in\mathcal{G}$, $\mu\in\mathfrak{T}_\theta$, i.e. μ is shift-ergodic with density $\theta\in(0,1)$. Then*

$$\int\mu(dz)\ \left|\mathbb{E}_z^{g}\mathfrak{x}_\zeta(t)-\theta\right|^2 \underset{t\to\infty}{\longrightarrow} 0, \qquad \zeta\in\Xi.$$

Proof. By the expectation formula (2.59), the claim follows from the *abstract L^2-ergodic theorem* in Fleischmann (1978), applied to the asymptotically uniformly distributed laws $p(t,\zeta,\cdot)$, $t\geq 0$ (for fixed ζ), according to Lemma 2.16. ■

The *weak law of large numbers* (1.24) in Theorem 4.b) is an immediate consequence of the following property:

Lemma 4.41. *For $g\in\mathcal{G}^0$ and $\mu\in\mathfrak{T}_\theta$ with $\theta\in(0,1)$,*

$$\mathbb{V}\mathrm{ar}_\mu^{g}\left(t^{-1}\int_0^t ds\ \mathfrak{x}_\xi(s)\right) \underset{t\to\infty}{\longrightarrow} 0, \qquad \xi\in\Xi. \tag{4.42}$$

Proof. 1°. By shift-invariance, we may set $\xi=\mathbf{0}$. Write γ_t for the random variable $t^{-1}\int_0^t ds\ \mathfrak{x}_{\mathbf{0}}(s)$, $t\geq o$. From (2.59) we know that $\mathbb{E}_\mu^{g}\gamma_t \equiv \theta$. Hence, it suffices to show that

$$(4.43)\qquad \mathbb{E}^{q}_{\mu}\gamma_t^2 = 2\,t^{-2}\int_0^t ds \int_s^t ds'\ \mathbb{E}^{q}_{\mu}\mathfrak{x}_0(s)\mathfrak{x}_0(s') \underset{t\to\infty}{\longrightarrow} \theta^2.$$

2°. Obviously, as $t\to\infty$ we may restrict the domains of integration to $1\le s\le t-1$ and $s+1\le s'\le t$. By the Markov property and (2.59),

$$(4.44)\qquad \mathbb{E}^{q}_{\mu}\mathfrak{x}_0(s)\mathfrak{x}_0(s') = \mathbb{E}^{q}_{\mu}\mathfrak{x}_0(s)\sum_{\zeta} p(s'-s,0,\zeta)\mathfrak{x}_{\zeta}(s).$$

Moreover, according to Lemma 2.58 (n=2) we have

$$(4.45)\qquad \mathbb{E}^{q}_{\mu}\mathfrak{x}_0(s)\mathfrak{x}_{\zeta}(s) = \int\mu(dz)\ a_{z,0}(s)a_{z,\zeta}(s) + \int_0^s dr \sum_{\xi} p(s-r,0,\xi)p(s-r,\zeta,\xi)\ \mathbb{E}^{q}_{\mu}g(\mathfrak{x}_{\xi}(r)).$$

By shift-invariance, in the latter expectation we use $\mathbb{E}^{q}_{\mu}g(\mathfrak{x}_{\xi}(r))\equiv\mathbb{E}^{q}_{\mu}g(\mathfrak{x}_0(r))$. Since the random walk is symmetric, we may apply Chapman-Kolmogorov three times to rewrite (4.44) as follows:

$$(4.46)\qquad \mathbb{E}^{q}_{\mu}\mathfrak{x}_0(s)\mathfrak{x}_0(s') = \int\mu(dz)\ a_{z,0}(s)a_{z,0}(s') + \int_0^s dr\ p(s+s'-2r,0,0)\ \mathbb{E}^{q}_{\mu}g(\mathfrak{x}_0(r)).$$

By (2.59) and Lemma 4.40,

$$(4.47)\qquad \int\mu(dz)\ a_{z,0}(s)a_{z,0}(s') \underset{t\to\infty}{\longrightarrow} \theta^2.$$

To complete the proof it suffices to show that

$$(4.48)\qquad \int_1^{t-1} ds \int_{s+1}^t ds' \int_0^s dr\ p(s+s'-2r,0,0)\ \mathbb{E}^{q}_{\mu}g(\mathfrak{x}_0(r)) = o(t^2) \quad \text{as } t\to\infty.$$

3°. We will prove (4.48) by considering several domains of r,s',s values in the integral. Let us first restrict the interior integral additionally to $r\le 1$, then it can be bounded from above by $\le$ const$/(s+s'-2)$ since by Lemma 2.20,

$$(4.49)\qquad p(t,0,0) \le \text{const}/t, \quad t\ge 1.$$

Consequently, this part of (4.48) leads to a negligible term of order $O(t)$. Thus, we may restrict our attention to the remaining case $1\le r\le s$.

4°. For fixed $\varepsilon\in(0,1/2)$, we may choose constants $b, b^{\varepsilon}>0$ such that $g^{\varepsilon}\le g\le b f$. Assume for the moment that for some constants $c,C>0$ independent of ε,

$$(4.50)\qquad \mathbb{E}^{q}_{\mu}g(\mathfrak{x}_0(r)) \le C(\varepsilon + e^{-cr} + 1/b^{\varepsilon}\log r), \quad 1\le r\le s.$$

Then for the remaining part of (4.48) substitute $s'=u+1$ and use (4.49) to get the bound

$$\le \text{const}\int ds\int du\int dr\ 1\{1\le r\le s\le u\le t-1\}\ (s+u-2r+1)^{-1}\ (\varepsilon + e^{-cr} + 1/b^{\varepsilon}\log r).$$

Change the order of integration to obtain

$$= \text{const} \int_1^{t-1} dr\ (\varepsilon + e^{-cr} + 1/b^{\varepsilon}\log r) \int_r^{t-1} ds \int_s^{t-1} du\ (s+u-2r+1)^{-1}.$$

The interior integral equals $\log(s+t-2r) - \log(s-r+1/2) - \log 2$. Using $L(x) := x\log x - x = \int \log x$, we can further integrate with respect to s to get

(4.51) $$[L(2t-2r-1) - L(t-r)] - [L(t-r-1/2) - L(1/2)] - (t-r-1)\log 2.$$

But
$$L(2t-2r-1) = 2(t-r-1/2)\log 2 + 2L(t-r-1/2),$$

and L is monotonously non-decreasing on $(1,+\infty)$. Hence, (4.51) can be bounded from above by $\le (t-r)\log 2 + \text{const}$. The latter additive constant leads to a term of order $O(\varepsilon t + 1 + t/b^{\varepsilon}\log t)$. Splitting the final integral gives

(4.52) $$\int_1^{(t-1)/2} dr\ (\varepsilon + e^{-cr} + 1/b^{\varepsilon}\log r)(t-r) \le \text{const}\ t\ (\varepsilon t + 1 + t/b^{\varepsilon}\log t)$$

by (4.31), and

$$\int_{(t-1)/2}^{t-1} dr\ (\varepsilon + e^{-cr} + 1/b^{\varepsilon}\log r)\ (t-r)$$
$$\le \text{const} \left(\varepsilon + \exp[-c(t-1)/2] + 1/\log((t-1)/2)\right) t^2,$$

which is of the same order as the r.h.s. of (4.52). Summarizing, the l.h.s. of (4.48) is of order $t(\varepsilon t + 1 + t/b^{\varepsilon}\log t)$ as $t\to\infty$, for each $\varepsilon\in(0,1/2)$, which gives the claim (4.48). Consequently, to complete the proof it remains to show (4.50).

5°. Recall that ε is fixed and $b^{\varepsilon}\ell^{\varepsilon} \le g \le b\ell$. Since $\ell \le \ell^{\varepsilon}+\varepsilon$, it suffices to consider $\mathbb{E}_{\mu}^{g}\ell^{\varepsilon}(\mathfrak{z}_0(r))$. By Lemma 4.6, except an ε-error we may pass to $\mathbb{E}_{\mu^{\varepsilon}}^{g}\ell^{\varepsilon}(\mathfrak{z}_0(r))$, $\mu^{\varepsilon}\in\mathfrak{P}$ with $\mu^{\varepsilon}([\varepsilon,1-\varepsilon]^{\Xi})=1$. Then we are able to apply Proposition 4.24 to arrive at the desired estimate (4.50). This finishes the proof. ■

Remark 4.53. If μ is even a *product law* and $g\ge b\ell$ for some $b>0$ then
$$\mathbb{V}\mathrm{ar}_{\mu}^{g}\left(t^{-1}\int_0^t ds\ \mathfrak{z}_0(s)\right) = O(1/\log t) \quad \text{as} \quad t\to\infty.$$
In fact, based on the previous proof but in the boundary case $\varepsilon=0$ (which can be admitted in the present case $g\ge b\ell$) one has only to deal with the error term related to the statement (4.47), which under a product law is given by $p(s+s',0,0)\ \mathbb{V}\mathrm{ar}_{\mu}\mathfrak{z}_0(0)$. Integrate and normalize this as in (4.43), and use (4.49) to get the claimed order of the error term. Actually one can expect that under these assumptions the fourth centered moments have a decay of order $1/(\log t)^2$ as $t\to\infty$, which then implies the *strong* law of large numbers. □

Proof of the oscillation property (1.25). Fix $\mu\in\mathcal{I}_{\theta}$ and $g\in\mathcal{G}^0$. Without loss of generality we assume that $\xi=0$. Suppose that $\mathbb{P}^{g}_{\mu}(\limsup_{t\to\infty} \mathfrak{x}_0(t) < 1) > 0$. Then

$$\mathbb{P}^{g}_{\mu}\left(\limsup_{t\to\infty} t^{-1} \int_0^t ds\, \mathfrak{x}_0(s) < 1\right) > 0,$$

and by (1.19) we get

$$\mathbb{P}^{g}_{\mu}\left(\limsup_{t\to\infty} t^{-1} \int_0^t ds\, \mathfrak{x}_0(s) < \theta/2\right) > 0$$

since $0<\theta<1$. But this contradicts the weak law of large numbers (1.24). The complementary statement follows by reflection. ∎

5. Thinned out Systems

Here we want to verify Theorem 1. The proof proceeds in two main steps. First by comparison results from Section 4 we reduce to the Fisher-Wright case, give the proof for this situation using duality, and then in the second step we show via coupling that it suffices to consider a product measure as initial distribution.

5.a) Reformulation and Reduction to the Fisher-Wright Case

Fix a law $\mu\in\mathfrak{T}_\theta$, a diffusion coefficient $g\in\mathcal{G}^0$, and $\alpha\geq 0$. Since all countably many components of the thinned out system ${}^{\alpha}\mathfrak{X}(t)$, $t>0$, take values in $[0,1]$, we may apply the *method of moments:* Take any collection of finitely many *different* labels $\xi(1),\dots,\xi(n)$, and integers $m(1),\dots,m(n) \geq 1$ and show that

$$\text{(5.1)}\quad \mathbb{E}^{g}_{\mu} \prod_{i=1}^{n} \left({}^{\alpha}\mathfrak{x}_{\xi(i)}(t)\right)^{m(i)} \underset{t\to\infty}{\longrightarrow} \mathcal{E} \prod_{i=1}^{n} {}^{\alpha}\mathfrak{x}_{\xi(i)}(\infty) = \mathcal{P}\left({}^{\alpha}\mathfrak{x}_{\xi(1)}(\infty)=\dots={}^{\alpha}\mathfrak{x}_{\xi(n)}(\infty)=1\right)$$

where

$${}^{\alpha}\mathfrak{X}(\infty) = \left({}^{\alpha}\mathfrak{x}_{\xi}(\infty)\right)_{\xi\in\Xi}, \qquad \mathcal{L}({}^{\alpha}\mathfrak{X}(\infty)) = \int F_{\theta,\alpha}(d\theta')\, \pi_{\tilde{\theta}'}, \qquad F_{\theta,\alpha} = \mathcal{L}(\tilde{Y}(\alpha))$$

with the product law $\pi_{\tilde{\theta}}$ and the time transformed Fisher-Wright diffusion $\tilde{Y}$ from Definition 1.14. Next we will rewrite both sides of (5.1).

1°. The r.h.s. of (5.1) equals

$$\text{(5.2)}\qquad \int F_{\theta,\alpha}(d\theta')\, (\theta')^n = E(\tilde{Y}(\alpha))^n = E(Y^{\theta}(-\log\alpha'))^n, \qquad \alpha' := \alpha\wedge 1.$$

Let Z^n denote the *pure non-linear death process* on $\{1,2,\dots\}$ with jumps from k to $k-1$ at rate $\binom{k}{2}$, $k\geq 2$, and starting in n. Via *duality* of the standard Fisher-Wright diffusions Y^{θ}, $0<\theta<1$, and Z^n, $n\geq 1$, that is

$$\text{(5.3)}\qquad E(Y^{\theta}(s))^n = E\theta^{Z^n(s)}, \qquad s\geq 0,$$

and the relation to the probability laws $p_{n,(\cdot)}(\gamma)$ of the scaling limit Proposition 3.13, it can be shown that $P(Z^n(s)=k) = p_{n,k}(\alpha')$ with $s=-\log\alpha'$; see

Cox and Griffeath (1986), p. 357 (recall also the Remark 3.15). Therefore the r.h.s. of (5.2) coincides with

$$\sum_{k=1}^{n} \theta^{k}\, p_{n,k}(\alpha'); \tag{5.4}$$

2^{o}. To rewrite the l.h.s. of (5.1), let $\zeta(1),\dots,\zeta(n)$ denote the "spaced" labels which correspond to $\xi(1),\dots,\xi(n)$ according to the notation (1.13), and set $\varphi := \varphi(\alpha t) := m(1)\delta^{\zeta(1)} + \dots + m(n)\delta^{\zeta(n)}$. Then the l.h.s. of (5.1) equals

$$\mathbb{E}^{g}_{\mu}(\mathfrak{X}(N^{t}))^{\varphi(\alpha t)} = \mathbb{E}_{\mu}\mathbb{E}^{g}_{\mathfrak{X}(0)}(\mathfrak{X}(N^{t}))^{\varphi(\alpha t)}. \tag{5.5}$$

3^{o}. We shall now analyze the r.h.s. in (5.5) above further. By Lemma 4.6 combined with Remark 4.9 and since (5.4) is continuous in θ, we may assume from now on in this subsection that the initial law μ is even concentrated on $[\varepsilon,1-\varepsilon]^{\Xi}$ with $\varepsilon\in(0,1/2)$. Choose $\mathfrak{b}^{\varepsilon},\mathfrak{b}^{0}>0$ such that

$$g^{\varepsilon} = \mathfrak{b}^{\varepsilon}\ell^{\varepsilon} \le g \le \mathfrak{b}^{0}\ell. \tag{5.6}$$

Having in mind the comparison Proposition 4.10 and the fact that the limit expression (5.4) does not depend on the diffusion coefficient g, we study now first (5.5) for systems with diffusion coefficient $\mathfrak{b}^{\varepsilon}\ell^{\varepsilon}$ where $\mathfrak{b}^{\varepsilon}>0$ and $\varepsilon\in[0,1/2)$, which covers both comparison cases. Apply (4.12) to pass in (5.5) to

$$\mathbb{E}_{\mu}\mathbb{E}^{\mathfrak{b}^{\varepsilon}}_{\mathbf{L}^{\varepsilon}\mathfrak{X}(0)}(\mathbf{H}^{\varepsilon}\mathfrak{X}(N^{t}))^{\varphi(\alpha t)} = \mathbb{E}_{\mu}\mathbb{E}^{\mathfrak{b}^{\varepsilon}}_{\mathbf{L}^{\varepsilon}\mathfrak{X}(0)}(\mathfrak{X}(N^{t}))^{\varphi(\alpha t)} + O(\varepsilon) \quad \text{as} \quad \varepsilon\to 0,$$

where $O(\varepsilon)$ is uniform in t by the boundedness of $\mathfrak{X}$ and since $m(1),\dots,m(n)$ is fixed. Using the *duality* Lemma 3.2 with $\mathbf{z}=\mathbf{L}^{\varepsilon}\mathfrak{X}(0)$ for the interacting Fisher-Wright system $\mathfrak{X}$ with diffusion constant $\mathfrak{b}^{\varepsilon}$, the latter expectation expression can be written as

$$\mathbb{E}_{\mu}\mathbf{E}_{\varphi(\alpha t)}(\mathbf{L}^{\varepsilon}\mathfrak{X}(0))^{\eta(N^{t})} = \mathbf{E}_{\varphi(\alpha t)}\mathbb{E}_{\mu}(\mathfrak{X}(0))^{\eta(N^{t})} + O(\varepsilon) \quad \text{as} \quad \varepsilon\to 0, \tag{5.7}$$

again uniformly in t.

4^{o}. The analysis of the r.h.s. of (5.7) leads us now back to coalescing random walks. Note that each pair of particles in $\varphi=\varphi(\alpha t)$ has hierarchical distance 0 or $\alpha t + O(1)$ as $t\to\infty$, depending on whether the particles have the same position or not. The *approximation* Proposition 3.6 tells us that as $t\to\infty$ we may replace the expectation at the r.h.s of (5.7) by

$$\tilde{\mathbf{E}}_{\varphi*}\mathbb{E}_{\mu}(\mathfrak{X}(0))^{\tilde{\eta}(N^{t})}, \qquad \varphi* = \delta^{\zeta(1)}+\dots+\delta^{\zeta(n)}. \tag{5.8}$$

5°. Assume for the moment that the initial law $\mu\in\mathfrak{I}_\theta$ is even a *product measure* $\mu^\theta := \prod_{\xi\in\Xi}\mu_\xi^\theta$ on $[0,1]^\Xi$ where $\int\mu_\xi^\theta(dr)\; r \equiv \theta\in(0,1)$. Then (5.8) can be written as

$$(5.9)\qquad = \tilde{\mathbf{E}}_{\varphi*}\prod_{\xi\in\Xi}\int\mu_0^\theta(dr_\xi)\; r_\xi^{\tilde{\eta}_\xi(N^t)} = \tilde{\mathbf{E}}_{\varphi*}\;\theta^{\|\tilde{\eta}(N^t)\|}.$$

Now we evaluate the r.h.s. of (5.9). If $\alpha=0$, then by Proposition 2.43 the latter expression converges to θ as $t\to\infty$, and we are done since $Y^\theta(\infty)$ has law $\theta\delta_1 + (1-\theta)\delta_0$. Suppose $\alpha>0$. Then $\varphi*$ belongs to $\tilde{\Phi}(n,\alpha t,1)$ (recall (3.12)) for all sufficiently large t, and the *scaling limit* Proposition 3.13 yields the convergence of (5.9) to (5.4) as claimed.

This limit expression (5.4) does not depend on $\mathfrak{b}^\varepsilon$ and $\varepsilon\in[0,1/2)$. Hence, for the boundary cases in (5.6) we get the same limit (5.4) as $t\to\infty$, except the $O(\varepsilon)$ error term in (5.7). But since we may let $\varepsilon\to 0$, this means that we verified the theorem for *general* $\mathfrak{g}\in\mathcal{G}^0$, of course, in the present case of a product initial measure.

6°. To complete the proof of Theorem 1, it remains to show that in calculating the limit of the Fisher-Wright expression (5.8) to actually consider product initial laws. This will follow from the *coupling Proposition* 5.11 below.

5.b) Successful Coupling of Fisher-Wright Systems

Fix initial laws $\mu,\nu\in\mathfrak{I}_\theta$, that is two (possibly different) homogeneous and ergodic laws μ and ν on $[0,1]^\Xi$, both with density θ. Consider the coupling $[\mathfrak{X},\mathfrak{Y}]$ with a joint initial law Γ which has marginals μ,ν according to the Definition 4.5, but in the *Fisher-Wright case* $\mathfrak{g}=\mathfrak{b}\mathfrak{f}$, $\mathfrak{b}>0$ fixed. This coupling is called *successful* if

$$(5.10)\qquad \mathcal{E}\left|\mathfrak{x}_0(t) - \mathfrak{y}_0(t)\right| \longrightarrow 0 \quad \text{as} \quad t\to\infty.$$

Proposition 5.11 **(successful coupling).** *Given* $\mu,\nu\in\mathfrak{I}_\theta$, *the coupling* $[\mathfrak{X},\mathfrak{Y}]$ *of Fisher-Wright systems with joint initial law* $\Gamma=\mu\times\nu$ *is successful.*

Remark 5.12. Coupling arguments have been proved useful in the theory of interacting systems, when the spatial ergodicity of a system is *preserved* under the evolution even in the limit as $t\to\infty$; see Liggett and Spitzer (1981), Greven (1991), Cox and Greven (1991b). Here we will even employ this technique in a context where the *preservation of ergodicity does not hold.* □

Since we are in the clustering regime, the proof of the previous proposition is non-standard and needs a preparatory lemma.

For notational simplification, write $[\mathbf{x},\mathbf{y}]$ for the initial configuration $[\mathfrak{X}(0),\mathfrak{Y}(0)]$ of the coupling $[\mathfrak{X},\mathfrak{Y}]$, and $\Delta_\xi(t) := \mathfrak{x}_\xi(t)-\mathfrak{y}_\xi(t)$, $\xi\in\Xi$, $t\geq 0$. For given $\xi,\mathbf{x},\mathbf{y},\mathbf{w}$ introduce

$$l_\xi^{\pm}(\mathbf{x},\mathbf{y},\mathbf{w}) := \limsup_{t\to\infty} t^{-1}\int_0^t ds\ (\Delta_\xi(s))^{\pm}, \tag{5.13}$$

respectively. Let P denote the law of $\mathbf{w}$ (the collection of Wiener processes).

Lemma 5.14. *$l_\xi^{\pm}(\mathbf{x},\mathbf{y},\mathbf{w})$ is $\mu\times\nu\times P$-a.s. constant in $\xi,\mathbf{x},\mathbf{y},\mathbf{w}$ and will be denoted by $l^{\pm}$, respectively.*

Proof. 1°. First we show by contradiction that

$$l_\xi^{+}(\mathbf{x},\mathbf{y},\mathbf{w}) \quad \text{is } \mu\times\nu\times P\text{-a.s. constant in } \xi. \tag{5.15}$$

Assume that

$$\mu\times\nu\times P\left(l_\xi^{+}(\mathbf{x},\mathbf{y},\mathbf{w}) - l_\zeta^{+}(\mathbf{x},\mathbf{y},\mathbf{w}) > \varepsilon\right) > 0, \quad \text{for some } \xi\neq\zeta \text{ and } 0<\varepsilon<1/2. \tag{5.16}$$

Then

$$\mu\times\nu\times P\left(\limsup_{t\to\infty} t^{-1}\int_0^t ds\ (\Delta_\xi^{+}(s) - \Delta_\zeta^{+}(s)) > \varepsilon\right) > 0. \tag{5.17}$$

Fix a δ satisfying $0<\delta<\varepsilon/2$. Set $A := \{\mathbf{z}\in[0,1]^{\Xi};\ \mathbf{z}_\xi,\mathbf{z}_\zeta<\delta \text{ or } \mathbf{z}_\xi,\mathbf{z}_\zeta>1-\delta\}$. Since in the Fisher-Wright case Condition 1.20 is satisfied, we may twice apply the a.s. statement (1.19) for $m=2$ to get

$$\lim_{t\to\infty} t^{-1}\int_0^t ds\ 1\{\mathfrak{X}(s),\mathfrak{Y}(s)\in A\} = 1, \qquad \mu\times\nu\times P\text{-a.s.} \tag{5.18}$$

Thus, in (5.17) we may additionally introduce the indicator $1\{\mathfrak{X}(s),\mathfrak{Y}(s)\in A\}$ without loosing the positivity of the probability. But $1\{\mathfrak{X}(s),\mathfrak{Y}(s)\in A\} = 1$ implies that $\Delta_\xi^{+}(s) - \Delta_\zeta^{+}(s) \leq 2\delta$, which leads to the contradiction $\varepsilon<2\delta$. Therefore (5.16) cannot be true. Thus, $l_\xi^{+}(\mathbf{x},\mathbf{y},\mathbf{w})$ is $\mu\times\nu\times P$-a.s. constant in ξ, and we write $l^{+}(\mathbf{x},\mathbf{y},\mathbf{w})$ for it.

2°. Now we want to show that $l^{+}(\mathbf{x},\mathbf{y},\mathbf{w})$ is constant in $(\mathbf{x},\mathbf{y},\mathbf{w})$. Denote by T_ξ the shift by ξ. We know by homogeneity of the model (see Definition 1.6) and by (5.15) that

$$l^{+}(T_\xi\mathbf{x},T_\xi\mathbf{y},T_\xi\mathbf{w}) = l_\xi^{+}(T_\xi\mathbf{x},T_\xi\mathbf{y},T_\xi\mathbf{w}) = l_0^{+}(\mathbf{x},\mathbf{y},\mathbf{w}) = l^{+}(\mathbf{x},\mathbf{y},\mathbf{w}), \quad \xi\in\Xi,\ \mu\times\nu\times P\text{-a.s.},$$

that is $l^+(\mathbf{x},\mathbf{y},\mathbf{w})$ is $\mu\times\nu\times P$-a.s. invariant with respect to the *simultaneous* shift of $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{w}$. From the shift ergodicity of $[\mathfrak{X}(0),\mathfrak{Y}(0)]$ under $\mu\times\nu$ and the i.i.d. assumption on $\mathbf{w}$ we conclude that the random field $(\mathbf{x}_\xi,\mathbf{y}_\xi,\mathbf{w}_\xi)_{\xi\in\Xi}$ under $\mu\times\nu\times P$ is shift ergodic. Hence $l^+(\mathbf{x},\mathbf{y},\mathbf{w})$ is $\mu\times\nu\times P$-a.s. constant in $[\mathbf{x},\mathbf{y},\mathbf{w}]$.

Summarizing, $l^+_\xi(\mathbf{x},\mathbf{y},\mathbf{w})$ is $\mu\times\nu\times P$-a.s. a constant, denoted by l^+. In the same way, $l^-_\xi(\mathbf{x},\mathbf{y},\mathbf{w})$ $\mu\times\nu\times P$-a.s. equals a constant l^-. This proves the lemma. ■

Proof of Proposition 5.11. 1°. It suffices to show $l^+l^-=0$. In fact if $[\mathfrak{X}(\infty),\mathfrak{Y}(\infty)]$ is distributed according to a weak limit point L of $\mathcal{L}([\mathfrak{X}(t),\mathfrak{Y}(t)])$ as $t\to\infty$ (where $[\mathfrak{X},\mathfrak{Y}]$ is the coupling with joint initial law $\Gamma=\mu\times\nu$), we get $L(\{0,1\}\times\{0,1\})=1$ by the ergodic theorem (1.11), thus, $l^+=0$ or $l^-=0$ implies $L(\mathfrak{X}(\infty)\le\mathfrak{Y}(\infty))=1$ or $L(\mathfrak{X}(\infty)\ge\mathfrak{Y}(\infty))=1$, respectively. But since we know that $\mathcal{E}(\mathfrak{X}_\xi(\infty)-\mathfrak{Y}_\xi(\infty))\equiv 0$ under L, we can conclude that in fact $L(\mathfrak{X}(\infty)=\mathfrak{Y}(\infty))=1$. This finally yields that the coupling is successful, and it really remains to show $l^+l^-=0$. The proof is by contradiction.

2°. Suppose $l^+l^->0$. Fix $\xi,\zeta\in\Xi$. From the almost sure statement (1.19) of Theorem 4.a) (note that we are in the Fisher-Wright case) we know that $\Delta_\xi(t)$ is "typically" close to 1, 0, or -1. But the hypothesis $l^+l^->0$ and the very definition (5.13) of l^+,l^- imply that it is actually close to both 1 and -1, more precisely, it *infinitely often oscillates* between values close to 1 and -1 as $t\to\infty$. Moreover, again by (1.19), $\Delta_\zeta(t)$ a.s. *oscillates synchronously* with $\Delta_\xi(t)$. This will be used to derive a contradiction for the behavior at time points s where $\Delta_\xi(s)$ and $\Delta_\zeta(s)$ have different sign.

3°. It is proved in Cox and Greven (1991b), assertion *(ii)* in the proof of Lemma 4, that

$$\int_0^\infty dt\ \mathcal{E}\Big(|\Delta_\xi(t)|\ 1\{\text{sign}(\Delta_\xi(t))\neq\text{sign}(\Delta_\zeta(t))\}\Big)<\infty \tag{5.19}$$

where by definition $\text{sign}(r)=1$ if $r>0$, and $=-1$ otherwise. For $0<\delta<1/2$ define the event

$$E_{t,\delta}:=\Big\{\text{sign}(\Delta_\xi(t))\neq\text{sign}(\Delta_\zeta(t)),\ |\Delta_\xi(t)|\ge\delta\Big\}.$$

If in (5.19) we restrict the expectation additionally to $E_{t,\delta}$ then we can estimate from below to conclude that $\int_0^\infty dt\ 1(E_{t,\delta}) < +\infty$ a.s. It is standard to show that this means $\sup\{t; 1(E_{t,\delta})=1\} < +\infty$ a.s. All we use for this is that we deal with diffusions with bounded drift and diffusion coefficient, so that the duration of a stay within $E_{t,\delta}$ is bounded from below, say by $c>0$, with positive probability (depending on c). We omit the standard but tedious details. Consequently, by monotonicity in δ, we have proved that

(5.20) $$\sup\{t\geq 0;\ 1(E_{t,\delta})=1\} < +\infty \quad \text{for all } 0<\delta<1/2 \text{ a.s.}$$

4°. Recalling step 2° of the proof, we want to show that *if* the two-dimensional process $(\Delta_\xi(t),\Delta_\zeta(t))$ oscillates between $(1,1)$ and $(-1,-1)$ synchronously as $t\to\infty$, then (5.20) cannot hold, so that we arrive at a contradiction. To that end we shall establish in step 5° below the following *crossing property* of $(\Delta_\xi(t),\Delta_\zeta(t))$: With positive probability we find random time points $s\to\infty$ with $\Delta_\xi(s) = +\delta$ but $\Delta_\zeta(s) < 0$. Then, obviously $\sup\{t\geq 0; 1(E_{t,\delta})=1\} = +\infty$ a.s. This contradiction to (5.20) shows that $l^+l^-=0$ and therefore then finishes the proof of Proposition 5.11.

5°. It remains to verify the crossing property formulated in step 4° above. For constants $h>0$ and $0 < \delta < 1/8$, denote by $F_{h,\delta}$ the following event: There exists a (random) sequence $s_1,s_2,\dots$ such that for $n\geq 1$,

(5.21) (i) $s_{n+1}-s_n > h,\quad \Delta_\xi(s_n) = -\delta,\quad \Delta_\zeta(s_n) \leq -\delta,$

(ii) $\Delta_\zeta(s_n+s) < 0,\quad 0\leq s\leq h,$

(iii) $\mathfrak{x}_\xi(s_n) \notin [1/2 - \delta,\ 1/2 + \delta],$

Set
$$\alpha := \mathcal{P}(F_{h,\delta}),\quad \beta := \mathcal{P}\left(\Delta_\xi(s_n+s) = \delta \text{ for some } s\in[0,h] \,\middle|\, F_{h,\delta}\right).$$
We want to show that $\alpha\beta > 0$ for h,δ sufficiently small, which proves the claimed crossing property.

We start by showing that $\alpha>0$. If in the definition of $F_{h,\delta}$ we drop the requirement (ii) and (iii), then we get an event of positive probability because of the infinitely many synchronous oscillations as $t\to\infty$ described in 2°. If we now add (ii) again, then the positivity of the probability is maintained, since the *continuous semimartingale* Δ_ξ (with respect to the family of

σ-fields $\sigma\{(\mathfrak{X}(s),\mathfrak{Y}(s))_{s\le t}\}$, $t\ge 0$,) has bounded local characteristics. Also, since $(\mathfrak{x}_\xi,\mathfrak{y}_\xi)$ is a semimartingale where the components of the martingale part have increasing processes with density (with respect to Lebesgue measure) bounded away from 0 if the components are not in $[1/2 - \delta,\ 1/2 + \delta]$ we can additionally impose (iii). The details are straightforward but tedious. (For a detailed exposition of such techniques, see the Appendix of Cox and Greven (1991b).) Altogether we see that we can choose h,δ such that $\alpha>0$.

We come to showing that $\beta>0$. Since Δ_ξ is a semimartingale where the bounded variation part is bounded in t for $t\in[s,s+h]$, it suffices to show that the martingale part M, say, can produce excursions in time h of size δ with positive probability. This martingale M is in law equal to a time transformed Brownian motion B, that is $M(t) \overset{\mathcal{L}}{=} B(\langle\!\langle M\rangle\!\rangle_t)$. For the Brownian motion B the excursions required have positive probability. It suffices therefore to show that the time transformation increases between s_n and s_n+h by at least say $\varepsilon>0$ with positive probability uniformly in n. But the quadratic variation process of the martingale part M is given by

$$b\int_0^t ds\ [\sqrt{\ell(\mathfrak{x}_\xi(s))} - \sqrt{\ell(\mathfrak{y}_\xi(s))}]^2,\quad t\ge 0. \tag{5.22}$$

The integrand can be 0 (recall (iii)) only if $\Delta_\xi(s)=0,+1,-1$. The local time of $\Delta_\xi(s)$ at $0,+1,-1$ is zero if the drift does not vanish at the same time, for a whole interval. This is not possible since *independent* Wiener processes act on all components.

This finishes the proof of Proposition 5.11, hence also the proof of Theorem 1. ∎

6. Block Averages

In this section we will prove Theorem 2. Let $g\in\mathcal{G}^0$ and $\mu\in\mathcal{T}_\theta$. By the shift invariance, without loss of generality we may assume that $\xi=\mathbf{0}$. Theorem 2 is equivalent to: For $m\ge 1$, $n(1),\dots,n(m)\ge 1$, and $0\le\alpha(1)<\dots<\alpha(m)$

$$\mathbb{E}^g_\mu \prod_{i=1}^m \left(\mathfrak{x}_{\mathbf{0},[\alpha(i)t]}(N^t)\right)^{n(i)} \underset{t\to\infty}{\longrightarrow} E\prod_{i=1}^m(\tilde{Y}(\alpha))^{n(i)}. \tag{6.1}$$

The proof of (6.1) will proceed in the following steps. We will reformulate both sides of (6.1), reduce to the Fisher-Wright case by comparison, use the duality with η, approximate by the coalescing random walks $\tilde{\eta}$, and finally apply the multi-scale limit proposition.

1°. First we want to rewrite the r.h.s. of (6.1) in terms of the quantities introduced in (3.27). Recall that $\tilde{Y}(\alpha) = Y^\theta(-\log\alpha'(i))$ by definition, where $\alpha' := \alpha \wedge 1$, and Y^θ is the standard Fisher-Wright diffusion starting in θ (which corresponds to the *right* "boundary values" of $\tilde{Y}$ at $\alpha \ge 1$). We will show (cf. [8], Lemma 4):

$$E\left(\prod_{i=1}^{m} (\tilde{Y}(\alpha))^{n(i)}\right) = \sum_{k=1}^{\infty} \theta^k \, p_{n(1),\dots,n(m);k}(\alpha'(1),\dots,\alpha'(m);1). \tag{6.2}$$

For $m=1$, this was already used in Subsection 5.a), see (5.3) and (5.4). Let $m>1$ and set $r_i := -\log\alpha'(i)$, $1 \le i \le m$. By conditioning Y^θ on the time interval $[0,r_2]$ and using the Markov property we obtain

$$E\left(\prod_{m\ge i\ge 1} (Y^\theta(r_i))^{n(i)}\right) = E\left(\left[\prod_{m\ge i\ge 2} (Y^\theta(r_i))^{n(i)}\right] E\left[(Y^\theta(r_1))^{n(1)} \,\Big|\, Y^\theta(r_2)\right]\right).$$

Applying $r_1 - r_2 = -\log(\alpha'(1)/\alpha'(2))$ and (6.2) for $m=1$ gives

$$= \sum_l p_{n(1);l}(\alpha'(1);\alpha'(2)) \, E\left(\prod_{m\ge i\ge 3} (Y^\theta(r_i))^{n(i)} (Y^\theta(r_2))^{l+n(2)}\right).$$

By induction and (an equivalent variant of) (3.27), we actually get (6.2).

2°. Using the definition (1.5) of block averages, we have to evaluate the limit of the l.h.s. of (6.1) in terms of products of certain $\mathfrak{x}_\zeta^{\mathcal{g}}(N^t)$. For this purpose, introduce again the index array $I := \{[i,u]; 1\le i\le m, 1\le u\le n(i)\}$, and natural numbers $k(i) \ge 1$, $1 \le i \le m$. In the various occurring sums it is useful to label components of the system with the help of some $\zeta(i,u) \in \Xi$, $[i,u] \in I$. For $s>0$ and $k(1),\dots,k(m) \ge 0$, we calculate

$$\mathbb{E}_\mu^{\mathcal{g}} \prod_{i=1}^{m} (\mathfrak{x}_{0,k(i)}(s))^{n(i)} = N^{-(\mathbf{k},\mathbf{n})} \, \mathbb{E}_\mu^{\mathcal{g}} \prod_{[i,u]\in I} \sum_{\|\zeta(i,u)\|\le k(i)} \mathfrak{x}_{\zeta(i,u)}(s) \tag{6.3}$$

where $(\mathbf{k},\mathbf{n})$ abbreviates the sum $k(1)n(1)+\dots+k(m)n(m)$. Continue with

$$= N^{-(\mathbf{k},\mathbf{n})} \sum_{\|\zeta(1,1)\|\le k(1)} \cdots \sum_{\|\zeta(m,n(m))\|\le k(m)} \mathbb{E}_\mu^{\mathcal{g}} \mathfrak{x}_{\zeta(1,1)}(s) \cdots \mathfrak{x}_{\zeta(m,n(m))}(s). \tag{6.4}$$

Specialize to $s=N^t$ and $k(i)=[\alpha(i)t]$.

3°. Here we will show that we may additionally restrict (again in a t-dependent way) the summation variables $\zeta(i,u)$ by the requirement

$$\|\zeta(i,u)-\zeta(j,v)\| \ge [\alpha(j)t] - \ell(\alpha(j)t), \qquad [i,u]\ne[j,v],\ i\le j, \tag{6.5}$$

(recall (2.25)). For $r \geq 1$ and elements $\zeta(i,u)$ and $\zeta(j,v)$ of the hierarchical group Ξ we have $\|\zeta(i,u)-\zeta(j,v)\| < r$ if and only if the first $r-1$ coordinates of both $\zeta(i,u)$ and $\zeta(j,v)$ are arbitrary but the remaining coordinates coincide. Hence, for fixed $[i,u] \neq [j,v]$ with $i \leq j$, the number of $\{\zeta(i,u),\zeta(j,v)\}$ where the inequality in (6.5) is violated, is bounded by $N^{[\alpha(i)t]+[\alpha(j)t]-\ell(\alpha(j)t)}$, since we may restrict our considerations to those t, where $[\alpha(i)t] < [\alpha(j)t] - \ell(\alpha(j)t)$, whenever $i<j$. Each remaining $\zeta(k,w)$ runs as written in (6.4), i.e. it takes on $N^{[\alpha(k)t]}$ many values. Thus the total number of summands in (6.4) for which the condition (6.5) is violated is bounded by $C(\mathbf{n})\ N^{([\alpha t],\mathbf{n})}\ \varepsilon(t)$, where $C(\mathbf{n})$ is some constant, $([\alpha t],\mathbf{n})$ stands for $[\alpha(1)t]n(1)+\ldots+[\alpha(m)t]n(m)$, and $\varepsilon(t)$ abbreviates $\sum_{j=1}^{m} N^{-\ell(\alpha(j)t)}$. Note that $\varepsilon(t) \underset{t\to\infty}{\longrightarrow} 0$ except the case $\alpha(1)=0$.

But if $\alpha(1)=0$, then all $\zeta(1,u)$ are $\mathbf{0}$, and there is not any pair $[1,u]\neq[1,v]$ which violates (6.5). Then it is not necessary to include the term with $j=1$ into the definition of $\varepsilon(t)$, and $\varepsilon(t)\to 0$ also in this case.

The restriction of summation to those $\zeta(i,u)$, $[i,u]\in I$, satisfying (6.5) is justified because of the prefactor $N^{-([\alpha t],\mathbf{n})}$ in (6.4) and the boundedness by 1 of the expectation expression. Note that those $\zeta(i,u)$ also obey (3.26) (with $\alpha(j,t) \equiv \alpha(j)$ and $c=1$), since, using (2.28), $\|\zeta(i,u)-\zeta(j,v)\| \leq [\alpha(j)t]$.

4°. We finally come to an asymptotic evaluation of the l.h.s. of (6.1). Introduce again the notation $\varphi(i,t) := \delta^{\zeta(i,1)}+\ldots+\delta^{\zeta(i,n(i))}$, $1\leq i\leq m$, and $\varphi(t) := \varphi(1,t)+\ldots+\varphi(m,t)$. Denote with $A(t)$ the set of all those $\varphi(t)\in\Phi$ which correspond to the range of the $\zeta(i,u)$ in (6.4) but with the additional restriction (6.5). Then it remains to show that

$$N^{-([\alpha t],\mathbf{n})} \sum_{\varphi(t)\in A(t)} \mathbb{E}_{\mu}^{g}(\mathfrak{X}(N^{t}))^{\varphi(t)} \tag{6.6}$$

converges as $t\to\infty$ to the limiting term in (6.2).

5°. Since the limiting expression (6.2) is continuous in θ, as in step 3° in Subsection 5.a), we may reduce the further proof to the case of a Fisher-Wright diffusions $g=b\ell$. Moreover, by the coupling Proposition 5.11 we can pass to a product initial law $\mu=\mu^{\theta}\in\mathfrak{T}_{\theta}$. Then with the duality Lemma 3.2, we may rewrite (6.6) as

$$= N^{-([\alpha t],n)} \sum_{\varphi(t)\in A(t)} \mathbf{E}_{\varphi(t)} \mathbb{E}_{\mu}(\mathfrak{X}(0))^{\eta(N^t)}.$$

Since the number of particles in $\varphi(t)$ is fixed, and because together with (3.26) (with $\alpha(j)$ *independent* of t) also the assumptions of the approximation Proposition 3.6 are fulfilled, we can pass to the asymptotically equivalent expression

$$N^{-([\alpha t],n)} \sum_{\varphi(t)\in A(t)} \tilde{\mathbf{E}}_{\varphi^*(t)} \mathbb{E}_{\mu}(\mathfrak{X}(0))^{\tilde{\eta}(N^t)}. \tag{6.7}$$

Using the product measure assumption and $\mathbb{E}_{\mu}\mathfrak{x}_{\xi}(0) \equiv \theta$, we write (6.7) as

$$= N^{-([\alpha t],n)} \sum_{\varphi(t)\in A(t)} \tilde{\mathbf{E}}_{\varphi^*(t)} \theta^{\|\tilde{\eta}(N^t)\|}.$$

Note that $\varphi(t)=\varphi^*(t)$ for all sufficiently large t by (6.5), except the case $\alpha(1)=0$ where $n(1)\delta^0$ shrinks down to δ^0. In any case, $\varphi^*(t)$ fulfills the Assumption 3.25. Thus, by the multi-scale limit Proposition 3.28, the terms of the sum tend uniformly to the r.h.s. of (6.2), since for $\alpha(1)=0$ (the left hand side of) (6.2) does not depend on $n(1)$. This finishes the proof because of $N^{-([\alpha t],n)} \#A(t) \xrightarrow[t\to\infty]{} 1$. ■

7. Cluster Sizes

In this section we are going to provide the proof of Theorem 3 which by assumption is restricted to the Fisher-Wright case $g=b\ell$. In Subsection 7.a) we start with analyzing the appropriate properties of the dual system. The argument in 7.a) follows closely Bramson et al. (1986), whereas new arguments are needed in 7.b) and c). There we have to show that the size of blocks in a cluster can actually be transformed via the duality into a statement about coalescing random walks. Since the probabilities for $\{\mathfrak{x}_{\xi}\geq 1-\varepsilon\}$ or $\{\mathfrak{x}_{\xi}\leq\varepsilon\}$ are not directly obtainable from the moments, we have to work here a bit.

7.a) The Scaling Proposition 3.13 for Growing Solid Blocks

The following hypothesis is some analog of *The Proposition* in Bramson et al. (1986) on simple coalescing random walks on $\mathbb{Z}^2$. (The proof of it is beyond the scope of the present paper and will probably be discussed in the future in connection with tightness questions related to Theorem 2.) To formulate it, we introduce the notation

$$\psi(t) := \sum_{\xi:\|\xi\|\leq t} \delta^{\xi} \in \tilde{\Phi}, \qquad t>0, \tag{7.1}$$

describing the configuration that exactly the block of size [t] around the origin $\xi=0$ is filled up with particles (*"solid block"*). Recall that η denotes the coalescing random walk of Section 3.

Hypothesis 7.2. For each $\alpha\in(0,1)$,

$$\limsup_{t\to\infty} \mathbf{E}_{\psi(\alpha t)}\|\eta(N^t)\| < \infty,$$

$$\lim_{r\to\infty} \limsup_{t\to\infty} \mathbf{E}_{\psi(\alpha t)}\#\{\xi\in\Xi;\ \eta_\xi(N^t)>0,\ \|\xi\|\geq t+\log r\} = 0. \qquad \blacksquare$$

Roughly speaking, the $N^{[\alpha t]}$ initial particles "uniformly smeared" around 0 have at time N^t asymptotically only a finite mean number of survivors, and they are kept in the box of approximate size $t+\log r$. Recalling the $p_{n,k}$ from Subsection 3.d), this hypothesis will be shown to imply that the total number of particles at time N^t even has a well-defined limit distribution (for an explicit formula, see (1.3) in [2]):

Proposition 7.3 **(solid blocks scaling limit).** *For* $\alpha>0$,

$$\lim_{t\to\infty} \tilde{\mathbf{P}}_{\psi(\alpha t)}(\|\tilde{\eta}(N^t)\|=k) = \lim_{n\to\infty} p_{n,k}(\alpha\wedge 1) =: p_{\infty,k}(\alpha\wedge 1), \qquad k\geq 1.$$

Proof. First of all note that the limits $p_{\infty,k}(\gamma) := \lim_{n\to\infty} p_{n,k}(\gamma)$, $k\geq 1$, $0\leq\gamma\leq 1$, exist since $\sum_{k=1}^m p_{n,k}(\gamma)$, $1\leq k\leq m\leq n$, $0\leq\gamma\leq 1$, is monotone decreasing in n, namely by its probabilistic contents, see Proposition 3.13. The claimed identity will be proved by estimation.

For fixed $n\geq k\geq 1$ and sufficiently large t, by the definition (2.25) there exist t-dependent $\zeta^1,\dots,\zeta^n\in\Xi$ with $[\alpha t]-\ell(\alpha t) \leq \|\zeta^i\|, \|\zeta^i-\zeta^j\| \leq [\alpha t]$, $i\neq j$. Thus, by (3.12) we find $\varphi(t)\in\tilde{\Phi}(n,\alpha t,1)$ such that $\varphi(t)\leq\psi(\alpha t)$ for all sufficiently large t. Hence,

$$\limsup_{t\to\infty} \tilde{\mathbf{P}}_{\psi(\alpha t)}(\|\tilde{\eta}(N^t)\|\leq k) \leq \limsup_{t\to\infty} \tilde{\mathbf{P}}_{\varphi(t)}(\|\tilde{\eta}(N^t)\|\leq k) = \sum_{i=1}^k p_{n,i}(\alpha\wedge 1), \qquad n\geq k,$$

by Proposition 3.13. Letting $n\to\infty$, we conclude

$$\limsup_{t\to\infty} \tilde{\mathbf{P}}_{\psi(\alpha t)}(\|\tilde{\eta}(N^t)\|\leq k) \leq \sum_{i=1}^k p_{\infty,i}(\alpha\wedge 1), \qquad k\geq 1. \tag{7.4}$$

It remains to prove the "reversed" inequality

$$\liminf_{t\to\infty} \tilde{\mathbf{P}}_{\psi(\alpha t)}(\|\tilde{\eta}(N^t)\|\leq k) \geq \sum_{i=1}^k p_{\infty,i}(\alpha\wedge 1), \qquad k\geq 1. \tag{7.5}$$

Since $p_{\infty,i}(1)\equiv 0$, this is trivially fulfilled if $\alpha\geq 1$, and we may assume that $0<\alpha<1$. We will proceed similar as in Bramson et al. (1986): Let $M,r\geq 1$ be natural numbers, and choose $\alpha(1)$ and $\alpha(2)$ such that $\alpha < \alpha(1) < \alpha(2) < 1$. Important will be the event

(7.6) $$G_{M,r}(t) := \left\{\|\tilde\eta(N^{\alpha(1)t})\|\le M,\ \ \tilde\eta_\xi(N^{\alpha(1)t})=0 \ \text{ for } \ \|\xi\| \ge \alpha(1)t + \log r\right\}$$

which, roughly speaking, says that at time $N^{\alpha(1)t}$ we have only a bounded number of particles which moreover are not too much spread. Recalling the notation (3.12), use the Markov property at time $N^{\alpha(2)t}$ to estimate

$$\tilde{\mathbf P}_{\psi(\alpha t)}(\|\tilde\eta(N^t)\|\le k)$$
$$\ge \sum_{j=1}^{M} \sum_{\chi\in\tilde\Phi(j,\alpha(2)t,2)} \tilde{\mathbf P}_{\psi(\alpha t)}\Big(G_{M,r}(t) \cap \{\tilde\eta(N^{\alpha(2)t})=\chi\}\Big)\ \tilde{\mathbf P}_\chi\Big(\|\tilde\eta(N^t-N^{\alpha(2)t})\|\le k\Big).$$

From the scaling limit Proposition 3.13 we know that

$$\lim_{t\to\infty} \tilde{\mathbf P}_\chi\Big(\|\tilde\eta(N^t-N^{\alpha(2)t})\|\le k\Big) \ = \ \sum_{i=1}^{k\wedge j} p_{j,i}(\alpha(2)) \ \ge \ \min_{j\ge 1} \sum_{i=1}^{k\wedge j} p_{j,i}(\alpha(2)),$$

uniformly on $\tilde\Phi(j,\alpha(2)t,2)$. Hence,

$$\liminf_{t\to\infty} \tilde{\mathbf P}_{\psi(\alpha t)}(\|\tilde\eta(N^t)\|\le k) \ \ge \ \Big(\min_{j\ge 1} \sum_{i=1}^{k\wedge j} p_{j,i}(\alpha(2))\Big)\ \mathcal U(r,M),$$

where

(7.7) $$\mathcal U(r,M) := \liminf_{t\to\infty} \tilde{\mathbf P}_{\psi(\alpha t)}\Big(G_{M,r}(t) \cap \Big\{\tilde\eta(N^{\alpha(2)t}) \in \bigcup_{j=1}^\infty \tilde\Phi(j,\alpha(2)t,2)\Big\}\Big).$$

Since $\sum_{i=1}^{k\wedge j} p_{j,i}(\alpha(2))$ is non-increasing in j, the minimum on $j\ge 1$ will be realized by the limit $\sum_{i=1}^{k} p_{\infty,i}(\alpha(2))$ as $j\to\infty$. But by continuity, $p_{\infty,i}(\alpha(2))$ tends to $p_{\infty,i}(\alpha)$ as $\alpha(2)\downarrow\alpha$. Thus, to finish the proof of Proposition 7.3, it remains to show:

Lemma 7.8. *For $\alpha<\alpha(1)<\alpha(2)<1$ fixed,* $\liminf_{M,r\to\infty} \mathcal U(r,M) = 1$, *with $\mathcal U$ defined in* (7.7).

Proof. The first step is to note that, already for fixed M,r,

$$\limsup_{t\to\infty} \tilde{\mathbf P}_{\psi(\alpha t)}\Big(G_{M,r}(t) \cap \{\tilde\eta(N^{\alpha(2)t}) \notin \bigcup_{j=1}^\infty \tilde\Phi(j,\alpha(2)t,2)\}\Big) = 0.$$

In fact, apply the Markov property at time $N^{\alpha(1)t}$, and use the speed of spread Lemma 3.22 with $T(t)=r(t)=\alpha(1)t$, $s(t)=\alpha(2)t$, $\varphi(t)=\tilde\eta(N^{\alpha(1)t})$, and $c=1$. Hence, the main step of proof consists in showing that

$$\limsup_{M,r\to\infty} \limsup_{t\to\infty} \tilde{\mathbf P}_{\psi(\alpha t)}(\complement G_{M,r}(t)) = 0,$$

with $G_{M,r}$ defined in (7.6). To this end, it suffices to demonstrate that

(7.9) $$\limsup_{t\to\infty} \tilde{\mathbf E}_{\psi(\alpha t)}\|\tilde\eta(N^{\alpha(1)t})\| < \infty,$$

(7.10) $$\lim_{r\to\infty} \limsup_{t\to\infty} \tilde{\mathbf E}_{\psi(\alpha t)}\#\Big\{\xi\in\Xi;\ \tilde\eta_\xi(N^{\alpha(1)t})>0,\ \|\xi\|\ge\alpha(1)t+\log r\Big\} = 0.$$

But these properties follow from Hypothesis 7.2 above since $\alpha < \alpha(1)$ and $\eta\ge\tilde\eta$ by our coupling Convention 3.5. This finishes the proof of Lemma 7.8, hence also the proof of Proposition 7.3. ■

By dominated convergence (recall that $0<\theta<1$), as an implication of Proposition 7.3, for $\alpha>0$ we get

$$\text{(7.11)}\quad \lim_{t\to\infty}\left(\tilde{E}_{\psi(\alpha t)}\theta^{\|\tilde{\eta}(N^t)\|} + \tilde{E}_{\psi(\alpha t)}(1-\theta)^{\|\tilde{\eta}(N^t)\|}\right) = \sum_{k=1}^{\infty} p_{\infty,k}(\alpha\wedge 1)\ (\theta^k+(1-\theta)^k).$$

But

$$\text{(7.12)}\qquad \sum_{k=1}^{\infty} p_{\infty,k}(\alpha\wedge 1)\ (\theta^k+(1-\theta)^k) = P^{\theta}(T\geq\alpha), \qquad \alpha>0,\ 0<\theta<1,$$

where T was defined in (1.18). In fact, if $\alpha\geq 1$, use $p_{\infty,k}(1)\equiv 0$, which follows from $p_{n,n}(1)\equiv 1$, and $P^{\theta}(T\geq 1)=0$. On the other hand, if $0<\alpha<1$, see [8], before formula line (4.1) (recall the transformation $s\mapsto -\log s$ reverses order). Note that in this case $p_{\infty,(\bullet)}$ is a probability law on $\{1,2,\ldots\}$.

By (7.11) and (7.12), for a proof of Theorem 3 it remains to identify the l.h.s. in (7.11) as the limiting probability of the event

$$\text{(7.13)}\qquad \{S_t^{\varepsilon}\geq\alpha\} = \{\mathfrak{X}(N^t)\in\mathfrak{B}_{\alpha t}^{\varepsilon}\}, \qquad t,\alpha>0,\quad 0<\varepsilon<1/2,$$

with respect to $\mathbb{P}_{\mu}^{\mathfrak{b}}$ as first $t\to\infty$ and then $\varepsilon\to 0$. This will be done in the remaining two subsections.

7.b) Proof of Theorem 3 under Product Initial Laws

Here we restrict our considerations to the special case of a product initial measure $\mu=\mu^{\theta}\in\mathfrak{I}_{\theta}$. We will show next:

Lemma 7.14. *For fixed $0<q<1$, $\alpha>0$, and* $m\geq 1$,

$$\text{(7.15)}\qquad \lim_{t\to\infty}\mathbb{E}_{\mu}^{\mathfrak{b}}(\mathfrak{X}(N^t))^{m\psi(\alpha t)} = \lim_{t\to\infty}\tilde{E}_{\psi(\alpha t)}\theta^{\|\tilde{\eta}(N^t)\|}.$$

Proof. From the duality Lemma 3.2,

$$\text{(7.16)}\qquad \mathbb{E}_{\mu}^{\mathfrak{b}}(\mathfrak{X}(N^t))^{m\psi(\alpha t)} = E_{m\psi(\alpha t)}\mathbb{E}_{\mu}(\mathfrak{X}(0))^{\eta(N^t)}.$$

By our coupling Convention 3.5, $\eta(N^t)\geq\tilde{\eta}(N^t)$ if $\eta(0)\geq\tilde{\eta}(0)$, hence

$$\text{(7.17)}\qquad E_{m\psi(\alpha t)}\mathbb{E}_{\mu}(\mathfrak{X}(0))^{\eta(N^t)} \leq \tilde{E}_{\psi(\alpha t)}\mathbb{E}_{\mu}(\mathfrak{X}(0))^{\tilde{\eta}(N^t)} = \tilde{E}_{\psi(\alpha t)}\theta^{\|\tilde{\eta}(N^t)\|},$$

where we used that μ is a product measure μ^{θ} and $\mathbb{E}_{\mu}\mathfrak{X}_{\xi}(0)\equiv\theta$.

For an estimate in the opposite direction we may assume that $\alpha<1$. In fact, the r.h.s of (7.15) is part of the expression in (7.11) which by (7.12) disappears if $\alpha\geq 1$. Choose $\alpha<\alpha(1)<1$. From our Hypothesis 7.2 follows that for each $\delta>0$, with probability $\geq 1-\delta$ we have at most M particles at time

$N^{\alpha(1)t}$, and they are located in the block of size $\alpha(1)t + \log r$, $r \geq r_0$ around 0. By the approximation Proposition 3.6, the resulting state $\eta(N^t)$ at time N^t belongs to $\tilde{\Phi}$, and we derived the r.h.s. of (7.17), restricted to an event of probability $\geq 1-\delta$. Since δ is arbitrary, the proof is finished. ■

For fixed $\alpha>0$, set

(7.18) $$Y(t) := (\mathfrak{X}(N^t))^{\psi(\alpha t)}, \quad Z(t) := (1-\mathfrak{X}(N^t))^{\psi(\alpha t)}, \quad t>0.$$

Since the r.h.s. $u := \lim_{t\to\infty} \tilde{\mathbf{E}}_{\psi(\alpha t)} \theta^{\|\tilde{\eta}(N^t)\|}$ of the identity (7.15) does not depend on m, we conclude that with respect to $\mathbb{P}^b_\mu$ (where $\mu=\mu^\theta$)

(7.19) $$\mathcal{L}(Y(t)) \underset{t\to\infty}{\Longrightarrow} u\delta_1 + (1-u)\delta_0.$$

Similarly, by reflection and symmetry

(7.20) $$\mathcal{L}(Z(t)) \underset{t\to\infty}{\Longrightarrow} v\delta_1 + (1-v)\delta_0 \quad \text{with} \quad v := \lim_{t\to\infty} \tilde{\mathbf{E}}_{\psi(\alpha t)}(1-\theta)^{\|\tilde{\eta}(N^t)\|}.$$

Next we show:

Lemma 7.21. *For fixed* $\alpha>0$, *we have* $\lim_{\varepsilon\to 0} \lim_{t\to\infty} \mathbb{P}^b_\mu\left(\sum_{\|\xi\|\leq\alpha t} \mathfrak{x}_\xi(N^t) \leq \varepsilon\right) = v.$

Proof. We will use the fact just derived that the weak limit law of $Z(t)$ as $t\to\infty$ has mass v at 1. By definition,

$$\log Z(t) = \sum_{\|\xi\|\leq\alpha t} \log(1-\mathfrak{x}_\xi(N^t)) \leq - \sum_{\|\xi\|\leq\alpha t} \mathfrak{x}_\xi(N^t),$$

so that

$$Z(t) \leq \exp\left[- \sum_{\|\xi\|\leq\alpha t} \mathfrak{x}_\xi(N^t)\right].$$

Hence, for $0<\delta<1$,

(7.22) $$\mathbb{P}^b_\mu(Z(t) \geq 1-\delta) \leq \mathbb{P}^b_\mu\left(\sum_{\|\xi\|\leq\alpha t} \mathfrak{x}_\xi(N^t) \leq \sigma_\delta\right)$$

with some constants $\sigma_\delta \underset{\delta\to 0}{\longrightarrow} 0+$. Letting first $t\to\infty$ and then $\delta\to 0$ yields

(7.23) $$v \leq \liminf_{\varepsilon\to 0} \liminf_{t\to\infty} \mathbb{P}^b_\mu\left(\sum_{\|\xi\|\leq\alpha t} \mathfrak{x}_\xi(N^t) \leq \varepsilon\right).$$

Conversely, to each $\varepsilon\in(0,1/2)$ choose a constant $c_\varepsilon>1$ such that

$$\log(1-r) \geq - c_\varepsilon r, \quad 0<r<\varepsilon, \quad \text{and} \quad c_\varepsilon \underset{\varepsilon\to 0}{\longrightarrow} 1.$$

Let $\sum_{\|\xi\|\leq\alpha t} \mathfrak{x}_\xi(N^t) \leq \varepsilon$. Then even each term of the sum is $\leq\varepsilon$, and we can conclude that

$$\sum_{\|\xi\|\leq\alpha t} \log(1-\mathfrak{x}_\xi(N^t)) \geq - c_\varepsilon \sum_{\|\xi\|\leq\alpha t} \mathfrak{x}_\xi(N^t) \geq - c_\varepsilon \varepsilon.$$

Hence $Z(t) \geq \exp[-c_\varepsilon \varepsilon]$. Consequently,

$$\limsup_{t\to\infty} \mathbb{P}^b_\mu\left(\sum_{\|\xi\|\leq\alpha t} \mathfrak{x}_\xi(N^t) \leq \varepsilon\right) \leq \limsup_{t\to\infty} \mathbb{P}^b_\mu\left(Z(t) \geq \exp[-c_\varepsilon \varepsilon]\right) = v,$$

for each ε. Combined with (7.23), the claim follows. ■

By Lemma 7.21, reflection and symmetry,

$$\lim_{\varepsilon\to 0}\lim_{t\to\infty} \mathbb{P}^{\mathfrak{b}}_{\mu}\left(\Sigma_{\|\xi\|\le\alpha t}\ (1-\mathfrak{x}_{\xi}(N^t)) \le \varepsilon\right) = u. \tag{7.24}$$

Combined with Lemma 7.21,

$$\lim_{\varepsilon\to 0}\lim_{t\to\infty} \mathbb{P}^{\mathfrak{b}}_{\mu}\left(\mathfrak{x}(N^t)\in\mathcal{B}^{\varepsilon}_{\alpha t}\right) = u+v, \qquad \alpha>0.$$

Consequently, the limiting probability of the event (7.13) is given by (7.11) as we had to show. This completes the proof of Theorem 3 in the case of product initial laws.

7.c) Theorem 3 under General Initial Laws

Let $\mu\in\mathfrak{I}_{\theta}$ that is μ is shift-ergodic with density $\theta\in(0,1)$. The following lemma will be used for a reduction to a product measure case $\mu=\mu^{\theta}$.

Lemma 7.25. *For $\alpha>0$ and as $t\to\infty$,*

$$\tilde{\mathbf{E}}_{\psi(\alpha t)}\mathbb{E}_{\mu}(\mathfrak{x}(0))^{\tilde{\eta}(N^t)} \sim \tilde{\mathbf{E}}_{\psi(\alpha t)}\mathbb{E}_{\mu^{\theta}}(\mathfrak{x}(0))^{\tilde{\eta}(N^t)} = \tilde{\mathbf{E}}_{\psi(\alpha t)}\theta^{\|\tilde{\eta}(N^t)\|}.$$

Sketch of Proof. The idea of proof consists of applying the *ergodic theorem* mentioned already in the proof of Lemma 4.40, based on *asymptotically uniformly distributed* sequences $(\nu_n)_{n\ge 1}$ of probability laws on Ξ^m ($m\ge 1$ fixed). The latter means that the variational distance between ν_n and any shifted ν_n will disappear as $n\to\infty$:

$$\Sigma_{\xi\in\Xi^m}\ |\nu_n(\xi) - \nu_n(\xi+\zeta)| \underset{n\to\infty}{\longrightarrow} 0, \qquad \zeta\in\Xi^m.$$

From Proposition 7.3 we know that for $0<\alpha<1$,

$$\lim_{t\to\infty} \tilde{\mathbf{P}}_{\psi(\alpha t)}(\tilde{\eta}(N^t)\le k) \underset{k\to\infty}{\nearrow} 1,$$

i.e. roughly speaking, at time N^t we have only a bounded number of particles. Together with Lemma 2.26, for $k\ge 1$ this implies

$$\limsup_{t\to\infty} \tilde{\mathbf{P}}_{\psi(\alpha t)}\left(\tilde{\eta}_{\xi}(N^t)\tilde{\eta}_{\zeta}(N^t)>0 \text{ for some } \xi,\zeta\in\Xi \text{ with } \xi\ne\zeta \text{ and } \|\xi-\zeta\|\le k\right) = 0,$$

i.e. loosely speaking, the boundedly many points at time N^t spread away. Finally, the step distribution p of our random walk in Ξ is everywhere positive. Therefore, the spreading points are asymptotically uniformly distributed as $t\to\infty$.

Now μ is ergodic by assumption, hence the mentioned statistical ergodic theorem yields, roughly speaking, that $(\mathfrak{x}(0))^{\tilde{\eta}(N^t)}$ approaches $\theta^{\|\tilde{\eta}(N^t)\|}$ in mean square as $t\to\infty$. This results in the claim of the lemma in the case $0<\alpha<1$.

If $\alpha \geq 1$, the number of particles at time N^t additionally grows unboundedly, hence the first term in the lemma will disappear as $t \to \infty$, just as the other terms. ■

Completion of Proof of Theorem 3. It is easy to check that throughout the proof of Theorem 3 given in the last subsection we can replace the exact expression $\tilde{E}_{\psi(\alpha t)} \theta^{\|\tilde{\eta}(N^t)\|}$ in the product measure case by $\tilde{E}_{\psi(\alpha t)} \theta^{\|\tilde{\eta}(N^t)\|} + o(1)$, which finishes the proof for general initial measure. ■

REFERENCES

[1] BRAMSON, M., COX, J.T. and GREVEN, A. (1992). Ergodicity of critical spatial branching processes in low dimensions. *Ann. Probab.* (to appear).

[2] BRAMSON, M., COX, J.T., GRIFFEATH, D. (1986). Consolidation rates for two interacting systems in the plane. *Probab. Th. Rel. Fields* **73** 613-625.

[3] CHUNG, K.L. (1967). *Markov Chains with Stationary Transition Probabilities.* 2-nd. edition, Springer-Verlag, Berlin.

[4] COX, J.T. and GREVEN, A. (1991a). On the long term behavior of finite particle systems: a critical dimension example. In: *Random Walks, Brownian Motion and Interacting Particle Systems. A Festschrift in Honor of Frank Spitzer.* Eds. R. Durrett and H. Kesten. *Progress in Probab.* **28** 203-213. Birkhäuser, Boston.

[5] COX, J.T. and GREVEN, A. (1991b). Ergodic theorems for infinite systems of locally interacting diffusions. *Ann. Probab.* (to appear).

[6] COX, J.T. and GRIFFEATH, D. (1983). Occupation time limit theorem for the voter model. *Ann. Probab.* **11** 876-893.

[7] COX, J.T. and GRIFFEATH, D. (1985). Occupation times for critical branching Brownian motions. *Ann. Probab.* **13** 1108-1132.

[8] COX, J.T. and GRIFFEATH, D. (1986). Diffusive Clustering in the two dimensional voter model. *Ann. Probab.* **14** 347-370.

[9] COX, J.T. and GRIFFEATH, D. (1990). Mean field asymptotics for the planar stepping stone model. *Proc. London. Math. Soc.* **61** 189-208.

[10] DAWSON, D.A. and FLEISCHMANN, K. (1988). Strong clumping of critical space-time branching models in subcritical dimensions. *Stoch. Proc. Appl.* **30** 193-208.

[11] DAWSON, D.A. and GREVEN, A. (1992a). Hierarchical models of interacting diffusions: Multiple time scale phenomena, phase transition and pattern of clusterformation. *Probab. Theory Rel. Fields* (to appear).

[12] DAWSON, D.A. and GREVEN, A. (1992b). Multiple time scale analysis of hierarchically interacting systems. *Kallianpur Festschrift* (to appear).

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

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

[15] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications.* Vol. II, 2-nd edition, Wiley, New York.

[16] FLEISCHMANN, K. (1978). Mixing properties of cluster-invariant distributions. *Litovsk. Mat. Sb.* **18** 191-199.

[17] FLEISCHMANN, K. and GÄRTNER, J. (1986). Occupation time processes at a critical point. *Math. Nachr.* **125** 275-290.

[18] GREVEN, A. (1991). A phase transition for the coupled branching process, part 1: The ergodic theory in the range of finite second moments. *Probab. Th. Rel. Fields* **87** 417-458.

[19] ISCOE; I. (1986). Ergodic theory and a local occupation time measure-valued critical branching Brownian motion. *Stochastics* **18** 197-243.

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

[21] LIGGETT, T.M. and SPITZER, F. (1981). Ergodic theorems for coupled random walks and other systems with locally interacting components. *Z. Wahrsch. verw. Gebiete* **56** 443-468.

[22] NOTOHARA, M. and SHIGA, T. (1980). Convergence to genetically uniform state in stepping stone models of population genetics, *J. Math. Biol.* **10** 281-294.

[23] OHTA, T. and KIMURA, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genet. Res. Camb.* **22** 201-204.

[24] SAWYER, S. and FELSENSTEIN, J. (1983). Isolation by distance in a hierarchically clustered population. *J. Appl. Probab.* **20** 1-10.

[25] SHIGA, T. (1980a). An interacting system in population genetics. *J. Mat. Kyoto Univ.* **20** 213-242.

[26] SHIGA, T. (1980b). An interacting system in population genetics, II. *J. Mat. Kyoto Univ.* **20** 723-739.

[27] SHIGA, T. and SHIMIZU, A. (1980). Infinite-dimensional stochastic differential equations and their applications. *J. Mat. Kyoto Univ.* **20** 395-416.

[28] TAVARÉ, S. (1984). Line-of-descent and genealogical processes, and their applications in population genetics models. *Theoret. Population Biol.* **26** 119-164.

KLAUS FLEISCHMANN

INSTITUT FÜR ANGEWANDTE ANALYSIS UND STOCHASTIK

HAUSVOGTEIPLATZ 5-7

D-O-1086 BERLIN, GERMANY

fleischmann@iaas-berlin.dbp.de

ANDREAS GREVEN

INSTITUT FÜR MATHEMATISCHE STOCHASTIK

LOTZESTR. 13

D-W-3400 GÖTTINGEN, GERMANY

greven@namu01.gwdg.de

Veröffentlichungen des Instituts für Angewandte Analysis und Stochastik

Preprints 1992

1. D.A. Dawson and J. Gärtner: Multilevel large deviations.
2. H. Gajewski: On uniqueness of solutions to the drift–diffusion–model of semiconductor devices.
3. J. Fuhrmann: On the convergence of algebraically defined multigrid methods.
4. A. Bovier and J.-M. Ghez: Spectral properties of one–dimensional Schrödinger operators with potentials generated by substitutions.
5. D.A. Dawson and K. Fleischmann: A super–Brownian motion with a single point catalyst.
6. A. Bovier, V. Gayrard: The thermodynamics of the Curie–Weiss model with random couplings.
7. W. Dahmen, S. Prößdorf, R. Schneider: Wavelet approximation methods for pseudodifferential equations I: stability and convergence.
8. A. Rathsfeld: Piecewise polynomial collocation for the double layer potential equation over polyhedral boundaries. Part I: The wedge, Part II: The cube.
9. G. Schmidt: Boundary element discretization of Poincaré–Steklov operators.
10. K. Fleischmann, F. I. Kaj: Large deviation probability for some rescaled superprocesses.
11. P. Mathé: Random approximation of finite sums.
12. C.J. van Duijn, P. Knabner: Flow and reactive transport in porous media induced by well injection: similarity solution.
13. G.B. Di Masi, E. Platen, W.J. Runggaldier: Hedging of options under discrete observation on assets with stochastic volatility.
14. J. Schmeling, R. Siegmund–Schultze: The singularity spectrum of self-affine fractals with a Bernoulli measure.
15. A. Koshelev: About some coercive inequalities for elementary elliptic and parabolic operators.
16. P.E. Kloeden, E. Platen, H. Schurz: Higher order approximate Markov chain filters.

17. H.M. Dietz, Y. Kutoyants: A minimum–distance estimator for diffusion processes with ergodic properties.

18. I. Schmelzer: Quantization and measurability in gauge theory and gravity.

19. A. Bovier V. Gayrard: Rigorous results on the thermodynamics of the dilute Hopfield model.

20. K. Gröger: Free energy estimates and asymptotic behaviour of reaction–diffusion processes.

21. E. Platen (ed.): Proceedings of the 1^{st} workshop on stochastic numerics.

22. S. Prößdorf (ed.): International Symposium "Operator Equations and Numerical Analysis" September 28 – October 2, 1992 Gosen (nearby Berlin).