Probab. Theory Relat. Fields 91, 467–506 (1992)

Probability Theory and Related Fields

Ergodic theorems for the multitype contact process

Claudia Neuhauser ⋆

Department of Mathematics, University of Southern California, DRB 155, Los Angeles, CA 90889-1113, USA

Received March 19, 1990; in revised form September 11, 1991

Summary. This paper studies a process involving competition of two types of particles (1 and 2) for the empty space (0). Each site of the lattice $\mathbf{Z}^d$ is therefore in one of three possible states: 0, 1, or 2. Particles of each type die with rate 1, while an empty site becomes occupied by a particle of type i with rate $\lambda_i \cdot$(proportion of neighbors of type i). The set of neighbors of a site x is of the form $\{y: \|x-y\| \leqq J\}$, for a positive integer J and a norm $\|\cdot\|$. Assuming there are only 0's and 1's present at the beginning, the process reduces to the contact process, with the critical rate of survival of 1's being λ_c. The basic problem we address is the existence of equilibria in which both types of particles coexist. Without loss of generality, one can restrict to the case $\lambda_2 \geqq \lambda_1 > \lambda_c$ and in this case we show:

(1) If $\lambda_2 > \lambda_1$, and the initial state is translation invariant and contains infinitely many 2's, then the 1's go away and the process approaches the invariant measure of the contact process with only 2's and 0's present,

(2) If $\lambda_2 = \lambda_1$, and $d \leqq 2$, then clustering occurs: starting from a translation invariant initial measure with no mass on all 0's, the process converges weakly to a convex combination of the two invariant measures obtained with only one type of particles present, and

(3) If $\lambda_2 = \lambda_1$, and $d \geqq 3$, then there is a one-parameter family of invariant measures including both types.

1 Introduction

The Multitype Contact Process is a Markov process in which the state at time t is ξ_t: $\mathbf{Z}^d \to \{0, 1, 2\}$. We say that a site is vacant if $\xi(x) = 0$, and it is occupied by a particle of type 1 (resp. 2) if $\xi(x) = 1$ (resp. 2). We formulate the evolution as follows:

⋆ Graduate fellow 1989–90, partially supported by the Army Research Office through the Mathematical Sciences Institute at Cornell University and NSF

Present address: Now at the Department of Mathematics, University of Southern California

(i) 1's and 2's each die (i.e., become 0) at rate 1.
(ii) 1's (resp. 2's) give birth to 1's (resp. 2's) at rate λ_1 (resp. λ_2).
(iii) If the birth occurs at x, the offspring is sent to a site chosen at random from $\{y: y-x \in \mathcal{N}\}$ where $\mathcal{N}$ is the set of neighbors of 0. We assume $\mathcal{N}$ to be of the from $\mathcal{N}=\{x: \|x\| \leqq J\}$ for some positive integer J where $\|\cdot\|$ is some arbitrary norm.
(iv) If $\xi_t(y)>0$ then the birth is suppressed.

We think of this process as a model for a biological population of two species where the species may compete only over vacant sites.

We begin with some simple observations. First if only one type of particle is present, the system reduces to the basic d-dimensional contact process with neighborhood set $\mathcal{N}$ (see e.g., Liggett (1985) or Durrett (1988)). Let λ_c be the critical value of the contact process, i.e.,

$$\lambda_c = \inf\{\lambda: P(|\xi_t^0|>0 \text{ for all } t)>0\}$$

where $|\xi_t^0|$ denotes the number of occupied sites in the contact process starting with a single particle at 0. Throughout the paper we will assume $\lambda_2 \geqq \lambda_1 > \lambda_c$. For it is easy to see that if $\lambda_1 \leqq \lambda_c$, the 1's die out and we end up with the d-dimensional contact process for the 2's.

The point of this paper is to prove ergodic theorems for this model. Since we have two types of particles, it is natural to ask whether there are equilibria where both types of particles can coexist. The answer is sometimes yes and sometimes no depending on the following three cases: (i) $\lambda_2 > \lambda_1$, (ii) $\lambda_2 = \lambda_1$ and $d=1$ or 2, and (iii) $\lambda_2 = \lambda_1$ and $d \geqq 3$. In the first two cases we obtain no interesting behavior, only one type can survive. The interesting behavior shows up in the third case where we obtain a one-parameter family of extremal, stationary and translation invariant measures where the parameter ranges continuously from 0 to 1 and reflects the density of 1's in the limiting distribution.

Our first result shows that the case $\lambda_2 > \lambda_1$ is not interesting:

(1) Theorem. *If $\lambda_1 < \lambda_2$, the "1's die out". That is, if ξ_0 is translation invariant and $P(\xi_0(0)=2)>0$, then $\xi_t \Rightarrow \mu_2$ the limit starting from all sites occupied by particles of type 2.*

Here $\Rightarrow$ denotes weak convergence, which is in this setting just convergence of finite dimensional distributions.

We will explain the intuition behind this result after we state our results for the case $\lambda_1 = \lambda_2$ and describe how they are proved.

To state the next result let $\mathcal{S}$ be the set of translation invariant measures and let μ_i be the limit starting from $\xi_0(x) \equiv i (i=1, 2)$. We use ξ_t^μ to denote the multitype contact process with initial distribution μ.

(2) Theorem. *If $\lambda_1 = \lambda_2$ and $d \leqq 2$, clustering occurs. More precisely,*

(i) *if $\xi_0 = \mu \in \mathcal{S}$ has no mass on $\xi_0(x) \equiv 0$*

$$\xi_t^\mu \Rightarrow \alpha \mu_1 + (1-\alpha) \mu_2$$

as $t \to \infty$.

(ii) *Furthermore, for any initial configuration* ξ_0

$$P(\xi_t(x)=1, \xi_t(y)=2) \to 0$$

as $t \to \infty$. *This holds for any* $x, y \in \mathbf{Z}^d$.

We will say something about α below. The next result shows that the interesting behavior occurs if $d \geqq 3$ and $\lambda_1 = \lambda_2$. Let $\mathscr{I}$ denote the set of stationary measures.

(3) Theorem. *Suppose* $\lambda_1 = \lambda_2$ *and* $d \geqq 3$.

(i) *Let* ξ_0^θ *be the initial distribution in which the coordinates* $\xi_0^\theta(x)$ *are independent and* $=1$ (resp. 2) *with probabilities* θ (resp. $1-\theta$). *As* $t \to \infty$

$$\xi_t^\theta \Rightarrow \nu_\theta.$$

(ii) *If* $\mu \in \mathscr{S}$ *is ergodic*

$$\xi_t^\mu \Rightarrow \nu_\theta \quad \text{for some } \theta \in [0, 1].$$

(iii) *The extremal translation invariant stationary distributions are* $(\mathscr{I} \cap \mathscr{S})_e = \{\nu_\theta : 0 \leqq \theta \leqq 1\}$.

The answers in (2) and (3) are similar to those for the voter model (Holley and Liggett 1975). It approaches total consensus in $d \leqq 2$. In $d \geqq 3$ differences of opinion may persist. These results are proved by using "duality". The duals in the voter model perform simple random walks. Two voters in this model have the same opinion if their duals hit, and may have different opinions if they do not hit. The difference in the behavior comes from the fact that random walks are recurrent in $d \leqq 2$, and transient in $d \geqq 3$.

As in the voter model, the key to our proofs will be duality, but the duals for the multitype process are more complicated. If $\lambda_1 = \lambda_2$ we define the dual process starting from a single site x much as in the contact process. There is a set of sites $\hat{\xi}_t^x$ so that if any of these sites are occupied then x will be occupied at time t. However now there is a hierarchy. If we imagine that all the sites at time 0 are occupied by particles of different color then there is one site that will paint x its color. We will denote this site by $\hat{\xi}_t^x(1)$ to indicate that this is the first member in the hierarchy. We call this ancestor the *distinguished particle*. Here and in what follows, *italics* indicate that we are giving a technical meaning to a phrase. If we make $\hat{\xi}_t^x(1)$ vacant then the color will change to that of some site $\hat{\xi}_t^x(2)$, the second member in the hierarchy, and so on. To determine the limiting behavior of the process starting from all sites occupied it is sufficient to keep track of the location of the distinguished particle $\hat{\xi}_t^x(1)$. To follow the evolution of the distinguished particle we use an idea of Kuczek (1989) to break the evolution at certain points which we call renewal points.

The good thing about the renewal points is that they define an embedded random walk for the distinguished particle. More precisely, the spatial and temporal displacement between two consecutive renewals from an i.i.d. family. We will also obtain exponential bounds on the spatial and temporal displacement which gives us control over the distinguished particle between consecutive renewals.

The picture of the dual we have in mind is now the following. The embedded random walk tells us where the distinguished particle is at the renewals. The exponential bound on the displacement gives us control over the location of the distinguished particle between consecutive renewals. It says that with high probability the distinguished particle will stay within a set linearly growing in time which we will call *triangle* for obvious reasons. Whenever a renewal occurs the next triangle starts at the bottom of the preceding one. Hence we obtain the picture of a chain of connected triangles where we can find our distinguished particle with high probability.

When we talk about *collision of triangles* we mean that the triangles of two different duals overlap. *Gluing two duals together* means that the two distinguished particles of two duals coalesced.

The idea behind the proof of (2) is the same as in the proof for the voter model, namely we rely on the recurrence of 1 and 2 dimensional random walks. The main technical problem is that as soon as the triangles collide, the embedded random walks are no longer independent. To overcome this we show there is a probability $\delta>0$ independent of the starting points so that with probability at least δ we can bring any two distinguished particles within a fixed finite distance K without a collision of their triangles. As soon as they are within distance K, it is easy to see that with positive probability we can glue the two duals together, i.e., make their distinguished particles coincide. Once this is done standard arguments take over to show the desired result.

The existence of the stationary distributions in $d \geqq 3$ is easier than the proof of (2). All we have to prove is that if the distance between x and y is large there is a positive probability that the two duals starting at x and y will not collide.

Two main ingredients are needed to prove (3). The first one is what we will call "convergence of trees". By this we mean the following. We start the dual at a certain site and follow the path of its distinguished particle until time t. Then we go back a fixed number of renewals and look at the tree growing out of this renewal point. We will prove that the trees have a limiting distribution by showing that the joint distribution of $(\xi_t^x(k)-\xi_t^x(1))$ converges. This together with the continuous mapping theorem shows that the one dimensional distributions converge.

For the convergence of the higher dimensional distributions we need the other ingredient which basically says that two dual processes either coalesce or get separated and are asymptotically independent.

We will find the extremal stationary and translation invariant measures by starting with any translation invariant and ergodic measure and proving that each one converges to a certain ν_θ where $\theta \in [0,1]$. The ν_θ's will be shown to be mutually singular. This follows from the weak law of large numbers for the Cesaro average of the 1's. By the multiparameter ergodic theorem, this quantity actually converges almost surely. An easy extension of this result will then show that they are ergodic. θ characterizes the density of the 1's in the limit.

It is fairly difficult to tell which ν_θ the system converges to except in the two cases where we either start with a translation invariant and ergodic measure without 0's in which case θ is the density of the 1's; or with a Bernoulli measure

μ in which case θ is characterized by the 1-density $\mu\{\xi: \xi(0)=1 \mid \xi(0)>0\}$. Likewise we cannot say much about the α in (2) except in the two cases just mentioned.

We want to note that the qualitative behavior of the system does not change if we have more than two species. If the birthrates are all the same, clustering will occur in $d \geqq 2$. In $d \geqq 3$ we will obtain a k-parameter family of extremal, translation invariant stationary distributions where k is the number of species-1. If the birthrates are different, then only the species with the highest birthrates will survive (starting from infinitely many species in each class). The proofs of the theorems rely heavily on the fact that both species have the same death rates. If particles of type i die at rate δ_i and give birth at rate λ_i, we conjecture that

$$\frac{\lambda_1}{\delta_1}=\frac{\lambda_2}{\delta_2}$$

is the right hypothesis for coexistence to occur in $d \geqq 3$ and for clustering to occur in $d \leqq 2$. In all other cases, the particle with the bigger quotient of λ_i/δ_i will eventually dominate. But we do not know how to prove this.

The paper is organized as follows: In Sect. 2 we give the graphical construction for the process and study the dual process. In this section we will also prove Theorem 1. The proof of this Theorem follows easily from the graphical construction and duality. This will be done at the end of Sect. 2. What we will basically show is that if we wait long enough, the renewal points in the dual process will be closed for the 1's. Section 3 is divided into three parts. In the first part we prove results that are also needed for the proof of Theorem 3. The second part proves Theorem 2 in dimension 1, and the third part proves it in dimension 2. We will relegate some preliminary results for the proof of Theorem 3 to Sect. 4. The proof of Theorem 3 will be carried out in Sect. 5. Formulas are numbered (1), (2), ... in each section. When formula (6) from Sect. 2 is referred to in a later section it is called (2.6).

2 Construction. Duality. Proof of Theorem 1

We begin by constructing the process from a collection of Poisson processes. The construction is basically the same as the one of the basic contact process (see e.g., Durrett 1988). We will first consider the case $\lambda_1=\lambda_2$. For x, $y \in \mathbf{Z}^d$ with $y-x \in \mathcal{N}$, let $\{T_n^{x,y}: n \geqq 1\}$ and $\{U_n^x: n \geqq 1\}$ be the arrival times of Poisson processes with rates $\lambda_2/|\mathcal{N}|$ and 1. At times $T_n^{x,y}$, we draw an arrow from x to y to indicate that if x is occupied then y will become occupied (if it is not already). At times U_n^x, we put a δ at x. The effect of a δ is to kill a particle at x (if it is present). An idea of Harris (1972) allows us to construct the process starting from any $\xi_0 \in \{0, 1, 2\}^{\mathbf{Z}^d}$. We will first construct the process up to time τ in such a way that $\mathbf{Z}^d$ splits into a countable number of a.s. finite components. Iterating this allows us to construct the process for all time.

We say that there is a path from $(x, 0)$ to (y, t) if there is a sequence of times $s_0=0<s_1<s_2<\ldots<s_n<s_{n+1}=t$ and spatial locations $x_0=x, x_1, \ldots, x_n=y$ so that:

(i) for $i=1, 2, \ldots, n$ there is an arrow from x_{i-1} to x_i at time s_i, and

(ii) the vertical segments $\{x_i\}\times(s_i, s_{i+1})$, $i=0, 1, \ldots, n$, do not contain any δ's.

We can now define an equivalence relation on $\mathbf{Z}^d\times[0,\infty)$ that defines the components. (x,t) and $(y,0)$ belong to the same equivalence class if (x,t) can be reached from $(y,0)$ (or (y,t) from $(x,0)$) by a path. By comparison with a branching process we can show that if we choose τ small enough such that the probability of a connection $<1/|\mathcal{N}|$, then all the components are a.s. finite. For more details see the paper cited above.

To take care of the case $\lambda_2>\lambda_1$, we start with the above construction for λ_2 and then toss a coin with success probability $(\lambda_2-\lambda_1)/\lambda_2$ at each arrow. If there is a success, we label the arrow with a "2" to indicate that only 2's can give birth through those arrows.

After constructing the process from the graphical representation we can define its dual process. If $\lambda_2>\lambda_1$, we define the dual process only for the 2's. If the rates are equal, it works for both types of particles. For the dual process we reverse the arrows and reverse time by mapping $\tilde{s}=t-s$. Let $\tilde{\xi}_t^x=\{y$: there is a path from $(x,\tilde{0})$ to $(y,\tilde{t})\}$. Since it is in general easier to work with a forward process than with a backward process, we will replace this process by the dual $\hat{\xi}_t^x$ that is constructed from a graphical representation that has arrows from x to y at rate λ_2 where y is again in the neighborhood set of x, and has δ's at rate 1. The construction can be done as before. We let

$$\hat{\xi}_t^x=\{y\text{: there is a path from }(x,0)\text{ to }(y,t)\}.$$

As in the basic contact process, $\hat{\xi}_t^x$ and $\tilde{\xi}_t^x$ have the same distribution. We will call the elements of $\hat{\xi}_t^x$ *ancestors* and the first ancestor sometimes *distinguished particle.*

As in the basic contact process the dual process tells us whether or not a site is occupied just by checking if at least one of the ancestors lands on an occupied site (if the process survives). Since the process we consider here has two types of particles, we are also interested in the type of particle sitting at a certain site. To figure this out note that the dual process $\{\hat{\xi}_s^x, 0\leqq s<t\}$ has a tree structure. We start the dual at $(x,0)$ and run it until time t where the ancestors land on the initial configuration. The tree structure defines an ancestor hierarchy in which the members are arranged according to the order they determine the type of the site $(x,0)$. We will describe the hierarchy now in greater detail by starting with the case $\lambda_1=\lambda_2$. Let $\hat{\xi}_t^x(n)$ be the nth member of the ordered ancestor set. If the first ancestor, $\hat{\xi}_t^x(1)$, lands on a 1 (resp. 2), then the site $(x,0)$ will be of type 1 (resp. 2). If it lands on a 0, we look at the type the second ancestor lands on, and so on. The first ancestor in the hierarchy that does not land on a 0, determines the type of the particle at $(x,0)$. If $\lambda_2>\lambda_1$, then a path that crosses an arrow labelled with a "2" is forbidden for a 1-particle. If $\hat{\xi}_t^x(1)$ does not cross any "2"-arrows and lands on an occupied site, then $\hat{\xi}_t^x(1)$ and $(x,0)$ are of the same type. If the first ancestor in the dual process lands on an empty site or crosses a "2"-arrow and lands on a 1-particle, we look at the second ancestor. If $\hat{\xi}_t^x(2)=1$ and the path that connects this site with $(x,0)$ does not cross any "2"-arrows, then the site $(x,0)$ will be occupied by a 1-particle. If $\hat{\xi}_t^x(2)=2$ and $\hat{\xi}_t^x(1)=0$, then the site $(x,0)$ will be occupied by a 2-particle. If $\hat{\xi}_t^x(2)=2$ and $\hat{\xi}_t^x(1)=1$, and if the first ancestor failed to paint $(x,0)$ its color, we check both particles for the first and second ancestor to determine whether or not the second ancestor can paint $(x,0)$ its color. For

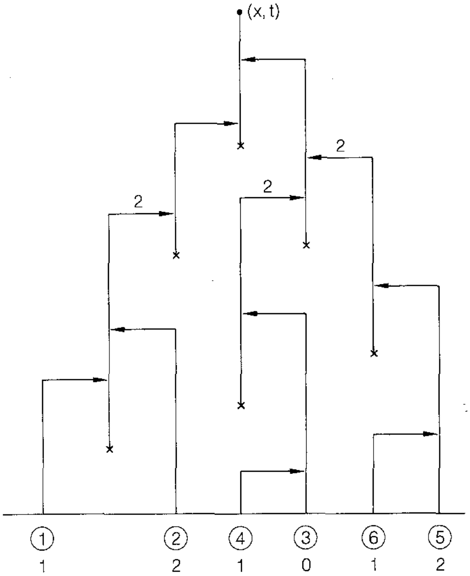

Fig. 1

instance, the second ancestor will fail if the first ancestor can still "use the path", i.e., has not crossed a 2-arrow by the time they both start using the same path. If the second ancestor fails too, we check the third ancestor in the 2-dual process, and so on. A picture is worth more than a hundred words, so we illustrate this at Fig. 1. The ordered set of the first six ancestors is $(x-3, x-1, x+1, x, x+3, x+2)$. If for instance, this ordered set is equal to (1, 2, 0, 1, 2, 1), then $(x, 0)$ will be of type 2 if $\lambda_2 > \lambda_1$, and of type 1 if $\lambda_2 = \lambda_1$ (then we ignore the "2's" at the arrows). In the first case, the fifth ancestor determines the type of $(x, 0)$; in the second case, it is the first ancestor.

Although the dual process looks more complicated than the one in the basic contact process, it has a nice property, which is the key to all of our proofs: We can break up the tree of paths at certain points into i.i.d. pieces and define an embedded random walk. We will call these points renewal points. This, together with estimates on how the tree behaves between the renewal points, allows us to trace the history of the tree by looking at the location of the renewal point.

For proving the announced property of the dual process we need some notation. We start the dual process at $(x, 0)$ and follow the path of the first ancestor. Whenever the first ancestor jumps to a site where it lives forever, we will call this site a renewal point. Since we are in the supercritical case, there is a positive probability that the dual process starting at x does not die out. We use a "restart argument" (see e.g., Durrett 1988, p. 72) to find a particle that lives forever: Pick a particle. If it does not live forever, wait until its family dies out and then pick another one. After at most a geometrically distributed number of trials we will find a family that lives forever and we can define the renewal points. Let $\Omega_{(x,0)}$ be the event that the dual process starting at x at time 0 lives forever. Most of the time we will suppress the dependence on x and 0 since by translation invariance the probability of this event does not depend on $(x, 0)$. We will then denote the event by Ω_∞. Let the spatial displacement between consecutive renewal points be X_i, and the corresponding temporal displacement be τ_i, then

$$S_n = x + \sum_{i=1}^{n} X_i \quad \text{and} \quad T_n = \sum_{i=1}^{n} \tau_i$$

will be the spatial and temporal location of the renewal point after the nth jump. This will also define the embedded random walk that jumps at random times T_n to the site $x + S_n$.

All this is summarized in the following Proposition which is a statement about the contact process. Its proof uses an idea of Kuczek (1989).

(1) Proposition. *Conditioned on the event* Ω_∞, $\{(X_i, \tau_i)\}_{i \geqq 1}$ *form an* i.i.d. *family of random vectors on* $\mathbf{Z}^d \times \mathbf{R}^+$. *The tail distributions of* X_i *and* τ_i *have exponential bounds,* i.e., *there are constants* $C, \gamma \in (0, \infty)$ *such that*

$$P(|X_i| > t) \leqq C e^{-\gamma t} \quad \textit{and} \quad P(\tau_i > t) \leqq C e^{-\gamma t}$$

holds.

We denote constants whose value are of no interest, by C, γ or alike. These values may change from line to line which will then be clear from the context.

In the proof we will not condition on Ω_∞ but instead use the recipe above to find a particle that lives forever. We are going to prove Proposition 1 in several steps. First, we will show that τ_1 can be bounded by a random variable with exponentially decaying tails. Let $x_0 = x$ and $\sigma_0 = \inf\{s > 0 : \xi_s^{(x,0)}(1)$ hits a $\delta\}$. Let x_1 be the location of the distinguished particle after $\xi_s^{(x,0)}(1)$ hits a δ. (Recall that we called the first ancestor also distinguished particle.) For $k \geqq 1$ we define a sequence of random variables $\{(x_k, \sigma_{k-1})\}_{k \geqq 1}$ in the following way. If $\sigma_{k-1} < \infty$ and $\xi_{\sigma_{k-1}}^{(x,0)} \neq \emptyset$, let x_k be the location of the distinguished particle at time σ_{k-1} (the distinguished particle jumps at time σ_{k-1} to a new site if $\xi_{\sigma_{k-1}}^{(x,0)} \neq \emptyset$). If $\xi_{\sigma_{k-1}}^{(x,0)} = \emptyset$, we let $x_k = x$ and start over again. If $\sigma_{k-1} = \infty$ we are done and (x_k, σ_{k-1}) is the renewal point we seek. Let $\sigma_k = \inf\{s > \sigma_{k-1} : \xi_s^{(x_k, \sigma_{k-1})} = \emptyset\}$ for $k \geqq 1$. σ_k is defined until it is equal to infinity. The superscript (x_k, σ_{k-1}) indicates where the distinguished particle jumped after the tree starting at (x_{k-1}, σ_{k-2}) died. Whenever we start with a new tree or at a point that lives forever, the first random time is defined in analogue to σ_0. It is always the first time the dual process hits a δ after starting from a site where either a new tree begins or the dual lives forever. Note that σ_0 is defined differently from $\{\sigma_k\}_{k \geqq 1}$ since $(x, 0)$ is the starting point. We would like to point out that once we hit a renewal point, only this branch matters: We never need to look at other branches that started before the renewal point. For an illustration see Fig. 2.

Define $A_1 = \{(x_1, \sigma_0)$ lives forever or $\xi_{\sigma_0}^{(x,0)} = \emptyset\}$, and $A_k = A_1^c \cap \ldots \cap A_{k-1}^c \cap \{(x_k, \sigma_{k-1})$ lives forever or $\xi_{\sigma_{k-1}}^{(x,0)} = \emptyset\}$ for $k \geqq 2$. Let N be the first k when A_k occurs. Then by the definition of N and $\{\sigma_k\}_{k \geqq 0}$, on Ω_∞

$$\tau_1 = \sigma_{N-1} = \sigma_0 + \sum_{k=1}^{N-1} (\sigma_k - \sigma_{k-1}).$$

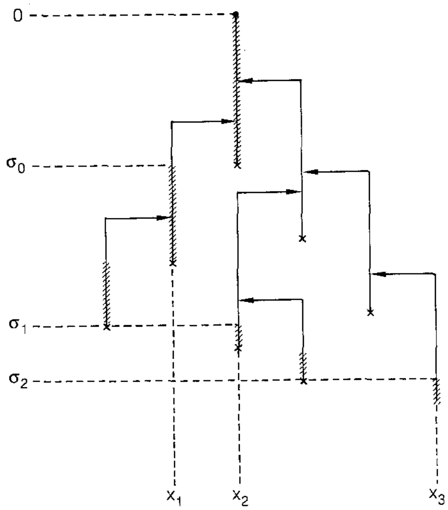

Fig. 2

We now want to show that the death of a branch happens early in the life of the branch. Bezuidenhout and Grimmett (1990) have shown for the basic contact process in d dimensions that if $\lambda_2 > \lambda_c$ then when viewed on suitable length and time scales the system dominates oriented percolation. Combining this with the analogue result for oriented percolation in Durrett (1984), Sect. 12, shows that there are constants C, $\gamma \in (0, \infty)$ so that

$$P(t < \sigma_k - \sigma_{k-1} < \infty) \leqq C e^{-\gamma t}.$$

The first in a series of lemmas that will show the announced property of τ_1 is

(2) Lemma. *On* $\{\sigma_k < \infty\}$

$$P(\sigma_{k+1} - \sigma_k \leqq m \mid \sigma_k, k < N) = P(\sigma_1 - \sigma_0 \leqq m).$$

Proof. The event $\{\sigma_{k+1} - \sigma_k \leqq m\}$ is determined by parts of the graph that are after σ_k and do not use any parts of the graph that are before σ_k; therefore $\{\sigma_k\}$ and $\{\sigma_{k+1} - \sigma_k \leqq m\}$ are independent. Since the graphical representation is translation invariant in time, the distribution of $\sigma_{k+1} - \sigma_k$ is the same as the one of $\sigma_1 - \sigma_0$. □

The next two lemmas will tell us that N and σ_{N-1} are a.s. finite.

(3) Lemma. *N is finite* a.s.

Proof. Using the definition of A_k and the formula for total probability we can write

$$\begin{aligned} P(N = n+1) &= P(A_1^c \cap \ldots \cap A_n^c \cap A_{n+1}) \\ &= P(A_{n+1} \mid A_1^c \cap \ldots A_n^c) \cdot P(A_n^c \mid A_1^c \cap \ldots A_{n-1}^c) \ldots P(A_2^c \mid A_1^c)\, P(A_1^c) \end{aligned}$$

By (2)

$$P(A_k^c | A_1^c \cap \dots A_{k-1}^c) = P(\sigma_1 - \sigma_0 < \infty; \xi_{\sigma_0}^{(x,0)} \neq \emptyset)$$

and

$$P(A_{n+1} | A_1^c \cap \dots A_n^c) = P(\sigma_1 - \sigma_0 = \infty \text{ or } \xi_{\sigma_0}^{(x,0)} = \emptyset).$$

Thus,

$$P(N = n+1) = P(\sigma_1 - \sigma_0 = \infty \text{ or } \xi_{\sigma_0}^{(x,0)} = \emptyset) \cdot P(\sigma_1 - \sigma_2 < \infty; \xi_{\sigma_0}^{(x,0)} \neq \emptyset)^n$$

from which it follows that $P(N < \infty) = 1$. □

(4) Lemma. $\sigma_{N-1} = \sigma_0 + \sum_{k=1}^{N-1} (\sigma_k - \sigma_{k-1})$ *is finite* a.s.

Proof. N is finite a.s. by Lemma (3), and given that $\{N = n+1\}$, $\sigma_1 - \sigma_0, \dots, \sigma_n - \sigma_{n-1}$ are finite with probability 1. From the definition of N it follows that σ_{N-1} is the first time after 0 that the particle lives forever or that the tree dies out. □

In the next step we will show that the $\{\sigma_k - \sigma_{k-1}\}_{k=1,\dots,n}$ are i.i.d. on $\{N = n+1\}$. We begin with

(5) Lemma.

$$P(\sigma_n - \sigma_{n-1} = t_n, \dots, \sigma_1 - \sigma_0 = t_1 | N = n+1) = \prod_{k=1}^{n} P(\sigma_1 - \sigma_0 = t_k | \sigma_1 - \sigma_0 < \infty; \xi_{\sigma_0}^{(x,0)} \neq \emptyset).$$

Proof. We write the left-hand side as

$$\frac{P(\sigma_n - \sigma_{n-1} = t_n, \dots, \sigma_1 - \sigma_0 = t_1, N = n+1)}{P(N = n+1)}.$$

The numerator can be written as

$$\begin{aligned} &P(\sigma_{n+1} - \sigma_n = \infty \text{ or } \xi_{\sigma_n}^{(x,0)} = \emptyset | \sigma_n - \sigma_{n-1} = t_n, \dots, \sigma_1 - \sigma_0 = t_1) \\ &\quad \cdot P(\sigma_n - \sigma_{n-1} = t_n, \dots, \sigma_1 - \sigma_0 = t_1) \\ &= P(\sigma_1 - \sigma_0 = \infty \text{ or } \xi_{\sigma_0}^{(x,0)} = \emptyset) \prod_{k=1}^{n} P(\sigma_1 - \sigma_0 = t_k; \xi_{\sigma_0}^{(x,0)} \neq \emptyset) \end{aligned}$$

by repeating the argument used in the proof of (2). Combining this with (3) proves the Lemma. □

We have to show that the distribution of τ_1 can be bounded by a random variable that has exponentially decaying tails. Let $g_{N-1}(s) = \mathbf{E} s^{N-1}$ be the moment generating function of the geometrically distributed random variable $N-1$. $g_{N-1}(s)$ can be extended beyond $s=1$. Let $\phi(\theta)$ be the moment generating function of $\sigma_1 - \sigma_0$ conditioned on $\sigma_1 - \sigma_0 < \infty$. Since the distribution of $\{\sigma_k - \sigma_{k-1}\}$ has exponentially bounded tails on $k < N$, we can find a $\theta_0 > 1$ so that

$$\mathbf{E}(e^{\theta \sum_{k=1}^{N-1} (\sigma_k - \sigma_{k-1})} | \sigma_1 - \sigma_0 < \infty) = g_{N-1}(\phi(\theta)) < \infty$$

for all $\theta \leqq \theta_0$. Hence $\mathbf{E} e^{\theta \tau_1} < \infty$ for all $\theta \leqq \theta_0$. This together with $e^{\theta t} P(\tau_1 > t) \leqq \mathbf{E} e^{\theta \tau_1}$ shows that there are constants $0 < C, \gamma < \infty$ so that

$$P(\tau_1 > t) \leqq C e^{-\gamma t}. \tag{6}$$

(6) and comparison with Richardson's model (see e.g., Durrett 1988, Sect. 1) tells us that the tree grows at most linearly in space. We therefore obtain a similar estimate for the spatial displacement between two consecutive renewals

$$P(|X_1| > t) \leqq C e^{-\gamma t}. \tag{7}$$

Note that since $P(\cdot | \Omega_\infty) \leqq P(\cdot)/P(\Omega_\infty)$ and $P(\Omega_\infty) > 0$, the exponential estimates also hold if we condition on survival. All that is left to show is that $\{(X_i, \tau_i)\}_{i \geqq 1}$ are independent and identically distributed on $\Omega_{(0,0)}$. Roughly speaking the family $\{(X_i, \tau_i)\}_{i \geqq 1}$ are independent since what happens before and after a certain renewal point uses disjoint parts of the graphical gadget and is therefore independent. Using translation invariance in time and space of the graphical gadget in time also shows that the family $\{(X_i, \tau_i)\}_{i \geqq 1}$ are identically distributed. To make the last two claims more precise, we will modify Kuczek's argument for the continuous time setting. We will define two quantities that will enable us to locate the points where the distinguished particle jumped to another branch of the tree.

Let $\xi_t^{(0,0)}(1)$ be the position of the distinguished particle at time t starting at $(0,0)$. Whenever $\xi_t^{(0,0)}(1)$ hits a "δ" it jumps to another site if the tree is still alive or we restart the process. The two quantities we are going to define are $\{U_k^{(0,0)}\}_{k \geqq 1}$ and $\{Y_k^{(0,0)}\}_{k \geqq 1}$. (We will suppress the superscript $(0,0)$ to save notation and define the quantities in the obvious way if we start at another location.) Let $U_0 = Y_0 = 0$. The first time the distinguished particle hits a δ and jumps to another site we set $U_1 = Y_1 = 1$. The next time it jumps we set $U_1 = 2$ and $Y_1 = 2$ if we can connect the new location with the one where we defined $U_1 = Y_1 = 1$. This will be called *jumping within a branch*. If not that is if we *leave a branch* we set $U_2 = 1$ and $Y_2 = 2$. We continue this way, i.e., whenever the distinguished particle jumps within a branch we increase U_i by 1 and set $Y_i = i + 1$. If the distinguished particle jumps to another branch we set $U_{i+1} = 1$ and $Y_{i+1} = Y_i$. We increase the superscript of Y_i only if the distinguished particle jumps to another branch of the tree. At the location of an (x_k, σ_{k-1}), $Y_k = k$. The next time it jumps we set $Y_k = k+1$ if it jumps within the branch and leave it that value as long as it stays within this branch. As soon as it jumps to another branch we set $Y_{k+1} = k+1$. With this algorithm $Y_k = k$ if (x_{k-1}, σ_{k-2}) cannot be connected to (x_k, σ_{k-1}) but any other location (x, t) to which the distinguished particle jumped in the meantime can be connected to (x_{k-1}, σ_{k-2}). $Y_k = k+1$ in all other cases.

U_k serves only as an auxiliary process to define the Y_k's. We use the Y_k's to locate the sites where the distinguished particle jumps to another branch. Now we can show

$$(8)\qquad \{(X_1,\tau_1)=(x_k,\sigma_{k-1})\}\cap\Omega_{(0,0)}=\{Y_k=k\}\cap\{\xi^{(0,0)}_{\sigma_{k-1}}(1)=x_k\}\cap\Omega_{(x_k,\sigma_{k-1})}.$$

To see this identity note that if $(0,0)$ is a renewal point, then $(0,0)$ can be connected by a path to (x_k,σ_{k-1}). If (x_k,σ_{k-1}) is the first renewal point after time 0, then $\{Y_k=k\}$ and $\Omega_{(x_k,\sigma_{k-1})}$. (If there was another renewal point before that, Y_k would be $k+1$.) On the other hand if (x_k,σ_{k-1}) is a renewal point and if $\xi^{(0,0)}_{k-1}(1)=x_k$, then $(0,0)$ can be connected to (x_k,σ_{k-1}) from which it follows that $(0,0)$ is a renewal point. Since $\{Y_k=k\}$, (x_k,σ_{k-1}) has to be the first renewal point. Using (8) it follows that

$$(9)\qquad P((X_1,\tau_1)=(x_k,\sigma_{k-1})|\Omega_{(0,0)})=\frac{P(Y_k=k;\xi^{(0,0)}_{\sigma_{k-1}}(1)=x_k;\Omega_{(x_k,\sigma_{k-1})})}{P(\Omega_{(0,0)})}$$
$$=\frac{P(Y_k=k;\xi^{(0,0)}_{\sigma_{k-1}}(1)=x_k)\cdot P(\Omega_{(x_k,\sigma_{k-1})})}{P(\Omega_{(0,0)})}$$
$$=P(Y_k=k;\xi^{(0,0)}_{\sigma_{k-1}}(1)=x_k).$$

In the second step we used that $\{Y_k=k\}\cap\{\xi^{(0)}_{\sigma_{k-1}}(1)=x_k\}$ only uses parts of the graph that are before (X_k,σ_{k-1}) whereas $\Omega_{(x_k,\sigma_{k-1})}$ uses parts that are after (x_k,σ_{k-1}). Hence the two events are independent. In the last step we used that $P(\Omega_{(0,0)})=P(\Omega_{(x_k,\sigma_{k-1})})=P(\Omega_\infty)$. Whenever we jump to a renewal point we relabel all the auxiliary quantities. To keep track of where we are we put a superscript on the quantities. For instance, $(X^{(l)}_k,\sigma^{(l)}_{k-1})$ is the location of the embedded process when it jumped k times after the lth renewal point. Then

$$\left\{\bigcap_{i=1}^{K}(X_i,\tau_i)=(x^{(i)}_{k_i},\sigma^{(i)}_{k_i-1})\right\}\cap\Omega_{(0,0)}$$
$$=\{Y_{k_1}=k_1\}\cap\{\xi^{(0,0)}_{\sigma_{k_1-1}}=x_{k_1}\}\cap\Omega_{(x_{k_1},\sigma_{k_1-1})}\cap\left\{\bigcap_{i=2}^{K}(X_i,\tau_i)=(x^{(i)}_{k_i},\sigma^{(i)}_{k_i-1})\right\}.$$

Using the same argument as before,

$$\{(X_2,\tau_2)=(x^{(2)}_{k_2},\sigma^{(2)}_{k_2-1})\}\cap\Omega_{(x_{k_1},\sigma_{k_1-1})}$$
$$=\{Y^{(x_{k_1},\sigma_{k_1-1})}_{k_2}=k_2\}\cap\{\xi^{(x_{k_1},\sigma_{k_1-1})}_{\sigma_{k_2-1}}(1)=x_{k_2}\}\cap\Omega_{(x_{k_2},\sigma_{k_2-1})}.$$

Repeating the argument and using translation invariance shows:

$$P\left(\left\{\bigcap_{i=1}^{K}(X_i,\tau_i)=(x_{k_i},\sigma_{k_i-1})\right\}\cap\Omega_{(0,0)}\right)=\prod_{i=1}^{K}P(Y_{k_i}=k_i;\xi^{(0,0)}_{\sigma_{k_i-1}}=x_{k_i})\,P(\Omega_{(0,0)})$$

Hence, conditioned on $\Omega_{(0,0)}$, $\{(X_i,\tau_i)\}_{i\geq 1}$ are i.i.d.

So far we haven't said anything about how the choice of the infection parameters λ_1,λ_2 affects the embedded random walk. In the case $\lambda_1=\lambda_2$ it is clear how to define the renewal points since all the arrows are unlabeled. If $\lambda_2>\lambda_1$ then we define the renewal points only for the 2's, i.e., starting the dual from say (x, t) we make use of all arrows to find the renewal points. Since the renewal points break the process into independent pieces, it is clear from the construction that with probability 1 there is a first time $T<\infty$ for which the renewal point is closed for the 1's in the sense that the 1's cannot pass through this renewal point. All the ancestors of this renewal point occupy positions in the ordered set of ancestors that are before the ancestors of earlier renewal points. Since the number of ancestors of a renewal point is roughly growing linearly in size, there will eventually be an ancestor occupied by a 2 in the first part of the ancestor vector that will succeed in painting x its color.

Proof of Theorem 1. Let $\lambda_2>\lambda_1$. We will show that

$$P(\xi_t(x)=1)\to 0$$

as $t\to\infty$ starting from a translation invariant initial configuration ξ_0.

We will now use the dual process $\tilde{\xi}_s^{(x,t)}$, $0\leqq s\leqq t$ which we start at (x, t) and determine the ordered set of ancestors after t units of time by going backwards in time on the graph. We will describe an algorithm that will find candidates in the dual process for painting x the color 2. Note that different ancestors can occupy the same site. But we can find arbitrarily many different candidates since the tree growing out of (x, t) is linearly growing in time. We will check the ancestors inductively.

In the first step we will find a subsequence in the set of ancestors that are candidates for painting x the color 2. In the second step we will extract a further subsequence so that all the candidates are different.

The first member of the subsequence is $\tilde{\xi}_t^{(x,t)}(1)$. We follow the path the distinguished particle takes to paint (x, t) its color until we first cross an arrow labelled with a 2. (Note: we are now going forward on the graph starting at $(\tilde{\xi}_t^{(x,t)}, 0)$.) Then we look backwards in time starting from the location where this particular arrow is attached. We discard all the offspring of this point. (Those are the next few members in the ancestor vector.) The first ancestor that is left after discarding those ancestors is the second member of the subsequence. We repeat the steps for this ancestor. We continue this until we run out of ancestors. We then extract a further subsequence so that all the candidates are different: We start with $\tilde{\xi}_t^{(x,t)}(1)$ and discard all members that occupy the same site as $\tilde{\xi}_t^{(x,t)}(1)$. Then we take the next ancestor that is left, and so on.

We denote the set of members of this subsequence by η_t. By choosing t large enough we can make the cardinality of η_t arbitrarily large. Denote by ζ_t the set of 2's in ξ_t. It suffices to show

$$P(\zeta_1\cap\eta_{t-1}=\emptyset)\to 0$$

as $t\to\infty$.

This will show that at least one of the candidates is occupied by a 2. The first one that is occupied by a 2 will paint x its color since by construction the 1's cannot go through. Given $\varepsilon>0$ and $M>0$ we can find $t>0$ so that

$$P(|\eta_{t-1}|<M)\leqq\varepsilon.$$

To finish the proof we will use Lemma 9.4 in Harris (1976) which in our context says that if ξ_0 is translation invariant with $P(\xi_0(0)=2)>0$ then given $\varepsilon>0$ there is an $M(\varepsilon)$ so that if $|\eta_{t-1}| \geqq M(\varepsilon)$ then

$$P(\eta_{t-1} \cap \zeta_1 = \emptyset) \leqq \varepsilon.$$

Therefore, if $\lambda_2 > \lambda_1$ and if the initial distribution is translation invariant, the 1's die out. This proves Theorem 1. □

3 Proof of Theorem 2

We will first describe the intuition behind the proof before we go into details. We will prove Theorem 2 in dimension 1 and 2 separately but the idea is basically the same for both dimensions. Only at one point we have to use a different method. From now on $\lambda_1 = \lambda_2 \equiv \lambda$.

We will show that we can bring two different distinguished particles starting at x and y with positive probability within a finite distance K. We do this in several steps. We define a rapidly increasing sequence of constants a_n where $a_{N_0} = K$ for some N_0, and $a_{N_1} = \|x-y\|$ for some N_1. We start the two duals at distance a_{N_1} and bring them within distance a_{N_1-1}, so that with high probability they behave independently throughout this time span. By doing the same step from starting at a_{N_1-1} and bringing them within distance a_{N_1-2}, and so on, we eventually get them within distance a_{N_0}. The estimates we obtain will hold uniformly in N_1. This iterating procedure has the advantage that we can use independence of the two dual processes as long as their triangles do not collide. As soon as they are within a finite distance K, there is a positive probability that we can glue the two particles forever together. What we do next is to show that both dual processes have renewals at the same time infinitely often with probability one where "same" is not to be taken too literally. This breaks the process down into independent pieces in the following way: We try to glue them together after they are brought within distance K; if we do not succeed, we wait until both of them have a renewal at the same time, then we start the whole iterating procedure again. By independence and Borel-Cantelli we will eventually succeed.

The difference in the proof for dimension 1 and 2 is in the part where we construct the single step for the iterating procedure. In dimension 1 we use Skorohod embedding to obtain the necessary estimates, in dimension 2 we use some potential theory for two dimensional random walks.

This section is organized as follows: In part a we will prove a series of lemmas that are valid in any dimension. Those lemmas are used to get the estimates for the iterating procedure. Part b is devoted to the proof of Theorem 2 in dimension 1, and part c for the proof in dimension 2.

a Preliminaries on the dual process

We will now state a series of lemmas that are needed in the proof of Theorems 2 and 3. Therefore we prove them in any dimension. Our first mission is to show that given two duals then the probability "they have renewals together", i.o. = 1. We will do this by showing that we can choose M such that in at least 1/3 of the time both duals have less than M particles. By killing all but one particle in each of the duals within a small amount of time, with positive probability both duals have their renewals together. By independence and Borel-Cantelli our result follows. Our second mission is to show that with positive probability we can glue two duals together that are within a finite distance K where we want one dual just being renewed and the other one having less than M particles. For proving the first result we start showing that the distribution of the number of particles in the tree since the last renewal has a limiting distribution. Let $\{(X_i, \tau_i)\}_{i \geqq 1}$ be the i.i.d. random vectors in $\mathbf{Z}^d \times \mathbf{R}^+$ defined in Sect. 2. Define the times and the spatial locations of the renewal points by $T_0 = 0$, $T_n = T_{n-1} + \tau_n$ and $S_0 = 0$, $S_n = S_{n-1} + X_n$.

(1) Lemma. *Let $v(s)$ = the number of particles in the tree at time s that have the last renewal point as an ancestor. Then there exist nonnegative numbers $p(k)$, $k = 1, 2, \ldots$ that sum up to 1 such that*

$$\frac{1}{t} \int_0^t \mathbf{1}_{\{v(s)=k\}} \, \mathrm{d}s \to p(k)$$

almost surely.

Proof. Let $N(t)$ be the number of jumps of the embedded random walk by time t. Then

$$\begin{aligned} (2) \qquad & \frac{1}{t} \sum_{l=1}^{N(t)} \int_0^{T_l - T_{l-1}} \mathbf{1}_{\{v(T_{l-1}+s)=k\}} \, \mathrm{d}s \\ & \leqq \frac{1}{t} \int_0^t \mathbf{1}_{\{v(s)=k\}} \, \mathrm{d}s \\ & \leqq \frac{1}{t} \sum_{l=1}^{N(t)+1} \int_0^{T_l - T_{l-1}} \mathbf{1}_{\{v(T_{l-1}+s)=k\}} \, \mathrm{d}s. \end{aligned}$$

Note that the sum is a sum of i.i.d. random variables. Let $m = \mathbf{E}\tau_1$. Since $\lim_{t\to\infty} \frac{N(t)}{t} \to \frac{1}{m}$ a.s. by the Renewal Theorem, we can apply the Strong Law of Large Numbers for a random number of summands (see e.g., Chung (1974), Chap. 5), and the left-hand side and the right-hand side of (2) converge a.s. to a constant we denote by $p(k)$. This shows (1). □

We can use (1) to show that we can choose M large enough such that in at least 1/3 of the time both duals have simultaneously less than M particles in the tree since the last renewal. We use the same notation as in (1) but put a subscript x on $v_x(s)$ to indicate that this quantity refers to the dual starting at x. Then

$$\begin{aligned} (3) \qquad & \frac{1}{t} \int_0^t \mathbf{1}_{\{v_x(s) \leqq M;\, v_y(s) \leqq M\}} \, \mathrm{d}s \\ & \geqq 1 - \frac{1}{t} \int_0^t \mathbf{1}_{\{v_x(s) > M\}} \, \mathrm{d}s - \frac{1}{t} \int_0^t \mathbf{1}_{\{v_y(s) > M\}} \, \mathrm{d}s \\ & \to 1 - 2 \sum_{l=M+1}^{\infty} p(l) \end{aligned}$$

where we used (1) and translation invariance in the last step. By choosing M such that $\sum_{l=1}^{M} p(l) \geqq 2/3$, the right-hand side of (3) is $\geqq 1/3$. This proves our claim.

We will call the dual process *good* at time t if the distinguished particle lives at time t and did not get any offspring since the last renewal. In the next Lemma we will consider two duals starting at different points and show that infinitely often both processes are good at the same time. This will also finish our first mission.

(4) Lemma. P (*both are good* i.o.) $=1$.

Proof. On the part of the state space where they coalesced we are done. Since then we have to deal with only one dual process. Otherwise, by (3) we can pick an M so that at least 1/3 of the time both of them have less than M particles. Hence by killing all particles but two at the right time we get

$$P(\text{both are good at the same time}) \geqq e^{-1} e^{-\lambda} [(1-e^{-1}) e^{-\lambda}]^{2M} \equiv \delta > 0.$$

Note that whenever both of them had a renewal at the same time we can start afresh and everything that happened after the renewals is independent from what happened before. The Borel-Cantelli Lemma now guarantees that both being good at the same time happens infinitely often with probability 1. Note that in the proof of this Lemma we do not assume that the duals are independent. □

Now we can turn to our second mission. We will show that there is a positive probability that both processes are good one unit of time after one of them had a renewal. This will be done in the following Lemma where we want to prove that with positive probability there is a time T where two dual processes that started at different sites and evolved independently until time T, are both good by time $T+1$. We will also assume that at time T one of the two processes has a renewal, and the other dual has less than $M+1$ particles in its tree since the last renewal. We consider two copies of the dual process starting at x and y. Let $I_T = \{\xi_t^x$ and ξ_t^y evolve independently for all $t \in [0, T]\}$. Denote by $J_T = \{(\xi_T^x(1), T)$ is a renewal point; $v_y(T) \leqq M\}$. With the sets just defined we can prove

(5) Lemma.

$$P(\textit{both processes are good by time } T+1 \mid I_T, J_T) \geqq e^{-(\lambda+1)(M+2)}$$

Proof.

$$\begin{aligned} &P(\text{both processes are good by time } T+1 \mid I_T, J_T) \\ &\geqq P(\text{all but one particle in } \xi_T^y \text{ die within one time unit and do not give birth}) \\ &\quad \cdot P(\xi_t^x(1) \text{ neither gives birth nor dies within } [T, T+1]) \\ &\geqq [(1-e^{-1}) e^{-\lambda}]^M [e^{-1} e^{-\lambda}]^2 \geqq e^{-(\lambda+1)(M+2)} \end{aligned}$$

by replacing $1-e^{-1}$ by e^{-1}. □

We will see later (when we use this Lemma) that we are by construction in the situation we conditioned on.

We will now show that as soon as the two distinguished particles are within distance K there is a positive probability that we can *glue* them *together*, i.e., they coalesce.

(6) Lemma. *Denote by A the event that two distinguished particles starting at distance K can be glued together on the set where both particles had a renewal at the beginning. Then there is a $\delta > 0$ so that*

$$P(A) \geqq \delta.$$

Proof. We will find a lower bound for the probability of the event A by estimating the following event. We require the particle located at y to survive for $3dK$ units of time without giving birth. Let $\|x-y\| \leqq K$. Then it takes the particle located at x at most dK steps to reach y. We will estimate this by observing that the probability of having a birth between time 1 and 2, and a death between time 2 and 3, can be bounded from below by $e^{-\lambda}(1-e^{-\lambda})e^{-2}(1-e^{-1})$. Note that it takes the particle at x at most $3dK$ units of time to hit the particle at y. The probability for the particle at y to survive until time $3dK$ without giving birth can be bounded from below by $e^{-(1+\lambda)3dK}$. Hence

$$P(A) \geqq e^{-(1+\lambda)3dK}[e^{-\lambda}(1-e^{-\lambda})e^{-2}(1-e^{-1})]^{dK} \equiv \delta > 0. \quad \square$$

b $d=1$

In this part we will show that with positive probability we can bring two duals within a finite distance K such that their triangles do not collide and as soon as they are within distance K, both are good. We have seen (Lemma 6) that two duals within distance K can be glued together with positive probability. If we do not succeed, we wait until both of them are good at the same time and do this procedure again. By independence and the Borel-Cantelli Lemma, this shows that eventually they coalesce with probability 1. This proves that the probability that two sites are different at time t goes to zero as $t \to \infty$. Hence, only one type can survive. Which type eventually survives depends on the initial measure.

Let $\{(X_i, \tau_i)\}_{i \geqq 1}$ be the i.i.d. random vectors in $\mathbf{Z} \times \mathbf{R}^+$ defined in Sect. 2 and let T_n and S_n be the times and the spatial locations of the renewal points as defined in Sect. 3a. We use Skorohod embedding (see e.g. Durrett (1990) or Billingsley (1986)) to embed S_n into a Brownian motion. S_n is the position of the random walk after the nth jump and it stays there put for $T_n \leqq t < T_{n+1}$, and jumps at time T_{n+1} to the location of the new renewal point. For the embedding define a Brownian motion B_t on $\mathbf{R}$. Let $\{U_n\}_{n \geqq 1}$ be a sequence of i.i.d. random variables with the same distribution as $|X_n|$, and let $\{\sigma_n\}_{n \geqq 1}$ be a sequence of stopping times defined by $\sigma_n = \inf\{t: |B_t - B_{T_n}| = U_n\}$. From this and the fact that the distribution of X_i is symmetric, it follows that B_{σ_n} has the same distribution as S_n. Furthermore, the $\sigma_1, \sigma_2-\sigma_1, \ldots$ are i.i.d. with $\mathbf{E}(\sigma_n-\sigma_{n-1})=\mathbf{E}X_1^2$ and $\mathbf{E}(\sigma_n-\sigma_{n-1})^2 \leqq 4\mathbf{E}X_1^4$.

All we need in the proof of Donsker's Theorem is that S_n and B_{σ_n} are defined on the same space and that $\frac{\sigma_n}{n} \xrightarrow{\text{a.s.}} \mathbf{E}X_1^2 \equiv \sigma^2 < \infty$ to obtain pathwise convergence, i.e., that

(7) $$\frac{S(n\cdot)}{\sigma\sqrt{n}} \Rightarrow B(\cdot).$$

We can certainly define S_n and B_{σ_n} on the same space; the second hypothesis follows from the fact that the tail distribution of X_1 has exponential bounds. Furthermore, by construction, we have

(8) $$|B_t - S_k| \leqq |X_{k+1}| \stackrel{d}{=} U_{k+1} \quad \text{for } T_k \leqq t < T_{k+1}.$$

Next we will prove that starting two dual processes at distance L, we can bring the two distinguished particles with positive probability within a distance of order $\log L$ and this uses up less than L^3 units of time.

In what follows we will make use of the fact that we can break up the dual process at the renewal points to obtain independent pieces. In between two consecutive renewal points the newborn particles stay with high probability within a set linearly growing in time. This was already shown by the exponential estimates in Sect. 2. If we start two dual processes at different sites, then they are independent as long as their triangles do not collide. This can be seen by defining the two dual processes $\hat{\xi}_t^x$ and $\hat{\xi}_t^y$ on two independent copies of the graphical representation G_1, G_2 (see Griffeath (1979), p. 21). We use G_1 for $\hat{\xi}_t^x$ and G_2 for $\hat{\xi}_t^y$. We define the embedded random walk $S_t^{(i)}$, $(i=1,2)$ on G_i in continuous time where the particle stays put between consecutive renewals and jumps to the site where the new distinguished particle lands at the renewal. As long as their triangles do not collide we let them evolve separately on the two gadgets. After a collision occurred we use only one copy, say G_1, to define the evolution of both dual processes.

We will now describe a single step in the iterative procedure that will bring the two distinguished particles close to each other. For this we will define a real valued random variable that measures the time it takes the two random walks to come within a certain distance: Let $\tilde{\tau}_L = \inf\left\{t : |S_t^{(1)} - S_t^{(2)}| \leqq \frac{12}{\gamma} \log L\right\}$ where $S_t^{(1)}$ and $S_t^{(2)}$ are the two embedded random walks corresponding to the dual processes $\hat{\xi}_t^0$ and $\hat{\xi}_t^L$, the one starting from $(0,0)$, the other from $(L,0)$. γ is the constant defined in (2.7). Note that in the proof of the following Lemma we do not use that $(0,0)$ or $(L,0)$ are renewal points. Recall that $N(t)$ denotes the number of jumps of the embedded random walk by time t. Let $m=\mathbf{E}\tau_1$. We will show

(9) Lemma. *There is a constant* $0<C<\infty$ *so that*

$$P\left(\max_{1\leqq i\leqq N(L^3)} |X_i| \leqq \frac{6}{\gamma}\log L \textit{ in both processes}; \tilde{\tau}_L \leqq L^3\right) \geqq 1 - \frac{C}{\sqrt{L}}$$

holds. γ *is the constant defined in* (2.7).

Proof. We will prove (9) in two steps. First we will show that

(10) $$P\left(\max_{1\leq i\leq N(L^3)}|X_i|>\frac{6}{\gamma}\log L\right)\leq\frac{A_2}{mL^3}$$

holds. We decompose the left-hand side of (10) according to whether $\{N(L^3)\leq 2L^3/m\}$ or $\{N(L^3)>2L^3/m\}$. This gives the following bound

$$\leq\frac{2L^3}{m}P\left(|X_1|>\frac{6}{\gamma}\log L\right)+P\left(N(L^3)>\frac{2L^3}{m}\right).$$

We use (2.7) to get a bound on the first term:

$$\frac{2L^3}{m}P\left(|X_1|>\frac{6}{\gamma}\log L\right)\leq\frac{2L^3}{m}A_3\,e^{-\gamma\frac{6}{\gamma}\log L}=\frac{2A_3}{mL^3}.$$

Since the τ_i form an i.i.d. sequence and have exponential bounds, we can use a large deviation estimate to obtain

(11) $$P\left(\frac{N(t)}{t}>\frac{2}{m}\right)\leq Ce^{-\gamma t}.$$

for some C, γ positive. With this the second term can be bounded by

$$P(N(L^3)>2L^3/m)\leq Ce^{-\gamma L^3}.$$

For the second part of the estimate note that by construction the two embedded random walks are independent until time $\tilde{\tau}_L$ on $\left\{\max_{1\leq i\leq N(L^3)}|X_i|\leq\frac{6}{\gamma}\log L \text{ in both processes}\right\}$. We will show

(12) $$P\left(\tilde{\tau}_L>L^3;\ \max_{1\leq i\leq N(L^3)}|X_i|\leq\frac{6}{\gamma}\log L \text{ in both processes}\right)\leq\frac{A_4}{\sqrt{L}}$$

We will be quite generous with this estimate. On the set where the triangles are small the two distinguished particles certainly come within distance $\frac{12}{\gamma}\log L$ if the difference of the embedded random walk is sufficiently negative, that is if their difference is $\leq -L$ for some $t\leq L^3$. We use the embedding of the random walk into Brownian motion to show (12). Let $\rho_x^{+1}=\inf\{t: B_t^1\leq x\}$ where B_t^1 is a Brownian motion starting in 1. Then

$$\begin{aligned}&P\left(\tilde{\tau}_L>L^3;\ \max_{1\leq i\leq N(L^3)}|X_i|\leq\frac{6}{\gamma}\log L \text{ in both processes}\right)\\&\leq P(\rho_{-1}^{+1}>L)=P(\rho_{-2}^{0}>L)=1-2P(B_L^0\geq 2)\\&=P\left(B_1^0\in\left[-\frac{2}{\sqrt{L}},\frac{2}{\sqrt{L}}\right]\right)\leq\frac{A_4}{\sqrt{L}}\end{aligned}$$

the third step coming from the reflection principle. This proves (12). Combining the two parts proves the Lemma. □

As we already mentioned in the introduction, the term *collision of triangles* denotes the event that the triangles of different duals overlap. Using the preceding lemmas we can prove

(13) Proposition. *Let two dual processes start within distance L. Then we can find a $K<\infty$ and a $\delta>0$ so that the following event has probability at least δ for all L: {it takes less than $2L^3$ time units to bring two duals within distance K. Their triangles do not collide and they are both good within one time unit after they came within distance K}.*

Proof. We now define the rapidly increasing sequence: $a_0=1$ and $a_n=e^{\frac{\gamma}{6}a_{n-1}}$.

Let $A_n=\left\{\max_{1\leqq i\leqq N(a_n^3)}|X_i|\leqq\frac{6}{\gamma}\log a_n \text{ for both processes; } \tilde{\tau}_{a_n}\leqq a_n^3\right\}$. Fix an η between 0 and 1. We can find a $N_0>0$ so that

$$C\sum_{k=N_0}^{\infty}\frac{1}{\sqrt{a_k}}\leqq 1-\eta \tag{14}$$

where C is the constant in (9). Set $K\equiv a_{N_0}$. Then

$$P\left(\bigcup_{k=N_0}^{\infty}A_k^c\right)\leqq C\sum_{k=N_0}^{\infty}\frac{1}{\sqrt{a_k}}\leqq 1-\eta. \tag{15}$$

Since $a_{N_1-1}^3+a_{N_1-2}^3+\ldots+a_{N_0-1}^3\leqq(N_1-N_0+2)\,a_{N_1-1}^3\leqq\exp(3\gamma a_{N_1-1}/6)=a_{N_1}^3$ if N_1 is large enough, we can also achieve that it takes up less than $2L^3$ time units to bring them within distance K by choosing N_0 sufficiently large. Note that as soon as they come within distance K, one of them has a renewal and the other process has less than M particles where M is of order K since the triangles in A_{N_0} are assumed to be of that size. We want both processes to be good within one time unit after they come within distance K, so we kill all but one particle in the process that did not have the renewal. (5) and (15) now give a bound on

$$\begin{aligned}&P\Bigg(\bigcap_{k=N_0}^{\infty}A_k;\text{both processes are good within one}\\&\qquad\text{time unit after they came within distance } K\Bigg)\\&\geqq\eta e^{-(\lambda+1)(M+2)}\equiv\delta>0\end{aligned}$$

since the two events are independent (they use nonoverlapping parts of the graphical gadget). This proves the Proposition. □

If we do not succeed in gluing the two duals together, we wait until both of them are good again and start the whole iterating procedure again. Everything is independent of what happened before and Borel-Cantelli shows that we can glue them eventually together with probability 1.

c $d=2$

So far we have proved Theorem 2 in dimension one. We turn now to the two dimensional case. It will turn out that the proof is very similar to the one for $d=1$. We will also use this iterating procedure in much the same way. Again, we need the exponential estimates proved in Sect. 2.

The other ingredient we will use, is an estimate on random walks. Let S_n be the position of the embedded random walk after the nth jump, as defined in Sect. 2. Again, the jumps take place at random times $\{T_n\}_{n\geqq 0}$, and the displacements are given by $\{X_n\}_{n\geqq 0}$. We need an estimate on how long it takes the random walk starting at $x\in\mathbf{Z}^2$ to hit a closed ball of radius r centered at 0. We denote the ball by $\bar{B}_r$. Let $\tau_r^x=\inf\{m: S_m\in\bar{B}_r|S_0=x\}$. Since $P(\tau_r^x>t)\leqq P(\tau_0^x>t)$, it is enough to find an estimate on the tail of the distribution of τ_0^x. By $P^m(x,0)$ we denote the probability that the random walk hits 0 in the mth jump starting at x. In the following we understand that all the summations range over integers.

$$\begin{aligned}\mathbf{E}_x[\#\text{ visits to 0 before } n] &= \sum_{m=0}^{n} P^m(x,0)\\ &= \sum_{m=0}^{n}\sum_{k=0}^{m} P(\tau_0^x=k)\,P^{m-k}(0,0)\\ &= \sum_{k=0}^{n} P(\tau_0^x=k)\sum_{m=0}^{n-k} P^m(0,0)\\ &\leqq P(\tau_0^x\leqq n)\left(\sum_{m=0}^{n} P^m(0,0)\right)\end{aligned}$$

Rearranging gives

$$P(\tau_0^x\leqq n)\geqq\frac{\sum_{m=0}^{n} P^m(x,0)}{\sum_{m=0}^{n} P^m(0,0)}.$$

Hence

$$P(\tau_0^x\geqq n)\leqq\frac{\sum_{m=0}^{n}[P^m(0,0)-P^m(x,0)]}{\sum_{m=0}^{n} P^m(0,0)}. \tag{16}$$

We will first find an estimate on the denominator $\sum_{m=0}^{n} P^m(0,0)$:

$$\sum_{m=0}^{n} P^m(0,0)\geqq C_1\int_{n_0}^{n}\frac{\mathrm{d}s}{s}\geqq C\log n \tag{17}$$

where we used the Local Central Limit Theorem for m large ($\geqq n_0$).

Let $a(x)=\lim_{n\to\infty}\sum_{m=0}^{n}[P^m(0,0)-P^m(x,0)]$. This is just the numerator of the right-hand side of (16) when $n=\infty$. The limit $a(x)$ exists (this is Theorem P1 on page 121 in Spitzer (1976)), and is called the potential kernel. Since $\mathbf{E}_0$(# visits to zero after $|x|^3)\approx\mathbf{E}_x$(# visits to zero after $|x|^3$), we can use an estimate on $a(x)$ to bound the numerator. Consulting page 124 in Spitzer (1976), we find Theorem P3 which we will state as

(18) Lemma. *A random walk satisfying*

(a) $P(x,y)$ *is two dimensional and aperiodic,*
(b) $\sum xP(0,x)=0$,
(c) $Q(\theta)=\sum(x\cdot\theta)^2 P(0,x)=\sigma^2|\theta|^2<\infty$,
(d) $\mathbf{E}[|X|^{2+\delta}]<\infty$ *for some* $\delta>0$, *has the property that*

$$\lim_{|x|\to\infty}\left[a(x)-\frac{1}{\pi\sigma^2}\log|x|\right]=C.$$

(a) clearly holds. For (b) note that our random walk is a difference of two copies of the embedded random walk, (d) holds since the distribution of the displacement has exponential bounds on its tails.

(c) is only true if the coordinates are i.i.d. which is not the case here. But looking again in Spitzer's book (Spitzer 1976) we can find on page 74 that for an irreducible, aperiodic random walk with mean vector 0 and second absolute moments finite, $Q(\theta)$ is positive definite, hence we can replace (c) by

(c′) $C_1|\theta|^2\leqq\sum(x\cdot\theta)^2 P(0,x)\leqq C_2|\theta|^2$

where $0<C_1\leqq C_2$ are the eigenvalues of the positive definite quadratic form $Q(\theta)$. Then we can still show that

$$\sum_{m=0}^{t}[P^m(0,0)-P^m(x,0)]\leqq C\log|x|$$

holds. The proof of the last statement can be easily adapted from Spitzer's book. So, we leave it to the reader. Returning to continuous time and combining (17) and (18) we get

$$P(\tau_0^x>t)\leqq C\frac{\log|x|}{\log|t|}. \tag{19}$$

Now we can define the blocks for the iterating procedure. Since it takes the random walk quite long to hit zero, we take a large time scale compared to the space scale. As in part b, denote by $\tau_{2^n}^{2^{n+1}}$ the time it takes the random walk to get within distance 2^n of the other random walk when they start within distance 2^{n+1}. Let

$$A_n=\{\max_{1\leqq i\leqq N(2^{(n+1)^3})}|X_i|\leqq\tfrac{1}{2}2^n \text{ for both processes};$$
$$\tau_{2^n}^{2^{n+1}}\leqq(2^{n+1})^{(n+1)^2}\}.$$

We will estimate $P(A_n^c)$ in the same way as in Lemma (9). From (19) we get

$$(20)\qquad P(\tau_{2^n}^{2^{n+1}} > (2^{n+1})^{(n+1)^2};\ \max_{1 \leq i \leq N(2^{(n+1)^3})} |X_i| \leq \tfrac{1}{2} 2^n \text{ for both processes})$$

$$\leq C \frac{\log 2^{n+1}}{\log(2^{n+1})^{(n+1)^2}} = \frac{C}{(n+1)^2}.$$

We also need an estimate on

$$P\left(N(2^{(n+1)^3}) > 2\frac{2^{(n+1)^3}}{m}\right) \leq C \exp\{-\gamma 2^{(n+1)^3}\}$$

where we used the large deviation estimate (11). This together with the exponential estimate (2.7)

$$P(|X_1| > \tfrac{1}{2} 2^n) \leq C e^{-\gamma 2^{n-1}}$$

can be used to estimate

$$P(A_n^c) \leq 2P(\max_{1 \leq i \leq N(2^{(n+1)^3})} |X_i| > \tfrac{1}{2} 2^n)$$

$$+ P(\tau_{2^n}^{2^{n+1}} > (2^{n+1})^{(n+1)^2};\ \max_{1 \leq i \leq N(2^{(n+1)^3})} |X_i| \leq \tfrac{1}{2} 2^n \text{ for both processes})$$

$$\leq 2\frac{(2^{n+1})^{(n+1)^2}}{m} C e^{-\gamma 2^{n-1}} + 2C \exp\{-\gamma 2^{(n+1)^3}\} + \frac{C}{(n+1)^2} \leq \frac{C}{(n+1)^2}$$

Now

$$P\left(\bigcup_{n=0}^{\infty} A_n^c\right) \leq \sum_{n=0}^{\infty} P(A_n^c) \leq \sum_{n=0}^{\infty} \frac{C}{(n+1)^2}.$$

The sum is finite. Again, as in the one dimensional case, pick N_0 so that

$$P\left(\bigcup_{n=N_0}^{\infty} A_n^c\right) \leq 1 - \eta$$

for a given η between 0 and 1. Then we can use the same technique as in the one dimensional case to conclude that

$$P\left(\bigcap_{n=N_0}^{\infty} A_n; \text{ both are good when they come within } 2^{N_0}\right) \geq \delta > 0.$$

Therefore we can bring them within distance $K = 2^{N_0}$ without collision of the triangles. The rest works as in the one dimensional case and we get that they coalesce with probability 1 as $t \to \infty$.

4 Preliminary results

In this section we will prove some preliminary results that we will need in Sect. 5 for the proof of Theorem 3. We will give a convergence determining class for the system and prove an estimate for the position of the distinguished particle that follows from the Local Central Limit Theorem for the embedded random walk. In addition, we will give an upper bound on the probability that two different distinguished particles will eventually coalesce.

Since the dual process of our system has some similarities with the dual process of the voter model, we can use ideas that were helpful there, to prove our results here. As we already said in Sect. 2, the distinguished particle does not quite perform a random walk but we can embed a random walk that allows us to keep control over the location of the distinguished particle. As in the case of independent random walks we can show that there is a positive probability that two distinguished particles will not hit each other if they start sufficiently far away from each other. From this we can conclude as in the voter model that both types of particles can survive.

We start with a description of a convergence determining class. It can be proved by inclusion-exclusion that any finite dimensional distribution can be written in terms of

$$\{\xi_t(x)=1 \text{ for all } x\in A;\ \xi_t(x)=2 \text{ for all } x\in B\}$$

where A and B are finite subsets of $\mathbf{Z}^d$. We omit the proof.

Before we can prove the next Lemma we have to introduce some notation. Let R_t be the location of the distinguished particle at time t and S_n be the location of the embedded random walk after the nth jump. By $N(t)$ we denote the number of jumps until time t. Let $T_n = T_{n-1} + \tau_n$ be the time of the nth jump (as in Sect. 2) and set $\mathbf{E}\tau_1 \equiv m < \infty$. Now we can prove

(1) Lemma. *If* $\frac{y_t}{\sqrt{t}} \to y$ *as* $t \to \infty$ *and* $\varepsilon > 0$, *then there is a constant* $C \in (0, \infty)$ *so that*

$$P(R_t = y_t) \leqq C t^{-d/2+\varepsilon}.$$

Proof. We want to show that

$$t^{d/2-\varepsilon} P(R_t = y_t) = \sum_n t^{d/2-\varepsilon} P(R_t = y_t; N(t) = n) \tag{2}$$

$$= \sum_n t^{d/2-\varepsilon} n^{-(d+1)/2} n^{(d+1)/2} P(R_t = y_t; N(t) = n)$$

is bounded by a constant $C \in (0, \infty)$. For this we will first find an estimate for $n^{(d+1)/2} P(R_t = y_t; N(t) = n)$. We can restrict ourselves to the case $n \in [t/m - t^{1/2+\varepsilon}; t/m + t^{1/2+\varepsilon}]$ for $\varepsilon > 0$ since $t^{d/2-\varepsilon} P(N(t) \notin [t/m - t^{1/2+\varepsilon}; t/m + t^{1/2+\varepsilon}]) \to 0$ as $t \to \infty$ by Chebyshev's inequality. We will denote the set $[t/m - t^{1/2+\varepsilon}; t/m$

$+t^{1/2+\varepsilon}]$ by A_t. We decompose $P(R_t=y_t; N(t)=n)$ according to the time of the last renewal and the location of the last renewal:

$$n^{(d+1)/2}P(R_t=y_t; N(t)=n)$$
$$=n^{(d+1)/2}\sum_z \int_0^t \mathrm{d}s P(S_n=y_t+z;\ T_n=t-s;$$
$$\tau_{n+1}>s;\ S_n-R_t=z).$$

Since T_n is a renewal time, $\{\tau_{n+1}>s; S_n-R_t=z\}$ and $\{S_n=y_t+z; T_n=t-s\}$ are independent. Hence we can factor the probability. Note that $\{\tau_{n+1}>s; S_n-R_t=z\}$ is independent of n, so we obtain

(3) $$=n^{(d+1)/2}\sum_z \int_0^t \mathrm{d}s P(\tau_1>s; R_s=z)$$
$$\cdot P(S_n=y_t+z;\ T_n=t-s).$$

Let $A_t=(-\alpha\log t, \alpha\log t)^d \cap \mathbf{Z}^d$. We will split the sum and the integral in (3) into

(4) $$n^{(d+1)/2}\left\{\sum_{z\in A_t}\int_0^{\beta\log t}\dots+\sum_{z\notin A_t}\int_0^{\beta\log t}\dots+\int_{\beta\log t}^t\dots\right\}.$$

The second and the third term do not contribute in the limit. So we will estimate them first. Throughout this and the next Section we denote by $\|\cdot\|$ the Euclidean norm for dimension $d>1$ and by $|\cdot|$ the Euclidean norm for $d=1$. The second term in (4) can be bounded by

$$\leqq n^{(d+1)/2}\sum_{z\notin A_t}\int_0^{\beta\log t}\mathrm{d}s P(R_s=z)$$
$$\leqq n^{(d+1)/2}\beta(\log t)\, P(\|X_1\|>\alpha\log t).$$

Using the exponential bound (2.7) on the triangles we get

(5) $$\leqq n^{(d+1)/2}\beta(\log t)\, C e^{-\gamma\alpha\log t}.$$

(2.6) gives an estimate on the third term in (4)

(6) $$\leqq n^{(d+1)/2}\int_{\beta\log t}^{\infty}\mathrm{d}s P(\tau_1>s)\leqq n^{(d+1)/2}\frac{C}{\gamma}e^{-\gamma\beta\log t}.$$

Summing (5) over $n\in A_t$ (the two additional terms coming from (2)) yields

(5′) $$\leqq\sum_{n\in A_t} t^{d/2-\varepsilon}n^{-(d+1)/2}n^{(d+1)/2}\beta(\log t)\, C e^{-\gamma\alpha\log t}$$
$$\leqq C t^{d/2-\varepsilon}\beta(\log t)\, t^{-\gamma\alpha}2t^{1/2+\varepsilon}=2C\beta(\log t)\, t^{(d+1)/2-\gamma\alpha}.$$

The right-hand side goes to 0 as $t \to \infty$ if α is chosen so that $(d+1)/2-\gamma\alpha=-1$. Summing (6) over $n \in A_t$ yields to

$$\text{(6')} \qquad \leqq \sum_{n \in A_t} t^{d/2-\varepsilon} n^{-(d+1)/2} n^{(d+1)/2} \frac{C}{\gamma} e^{-\gamma\beta \log t}$$

$$\leqq \frac{C}{\gamma} t^{d/2-\varepsilon-\gamma\beta} 2t^{1/2+\varepsilon} = \frac{C}{\gamma} t^{(d+1)/2-\gamma\beta}$$

which tends to 0 as $t \to \infty$ if β is chosen so that $(d+1)/2-\gamma\beta=-1$. This shows that the last two terms do not contribute in the limit $t \to \infty$.

The limiting behavior is determined by the first term. This time we have to estimate more carefully. We will approximate the integral over s in the first term in (4) by a sum in the following way: Decompose the interval $[0, \beta \log t]$ into disjoint intervals of length h: $[0, \beta \log t] = \bigcup_k (s_k + [0, h])$ where $s_k = kh$ and k runs over those integers for which $s_k \in [0, \beta \log t]$, i.e., $k=0, 1, \ldots, [(\beta \log t)/h]$. Note that $\{N(t)=n\} \subset \bigcup_k \{T_n \in t - s_k - [0, h];\ \tau_{n+1} > s_k\}$ where the union is over the same k's as above. With this we can estimate the first term in (4) by

$$\text{(7)} \qquad \leqq C \sum_{z \in A_t} \sum_{k=0}^{[(\beta \log t)/h]} P(\tau_1 > s_k; R_s = z)$$

$$\cdot n^{(d+1)/2} P(S_n = y_t + z;\ T_n \in t - s_k - [0, h]).$$

Since $(S_n; T_n) = \left(\sum_{i=1}^{n} X_i; \sum_{i=1}^{n} \tau_i \right)$ and the $\{(X_i; \tau_i)\}_{i \geqq 1}$ are i.i.d., it follows from the usual Local Central Limit Theorem with $h>0$ (see Stone 1967) that

$$\text{(8)} \qquad n^{(d+1)/2} P(S_n = u_n;\ T_n - nm \in v_n + [0, h]) \to h\Psi(u, v)$$

if $\frac{u_n}{\sqrt{n}} \to u$ and $\frac{v_n}{\sqrt{n}} \to v$. $\Psi(\cdot,\cdot)$ is proportional to a density of a bivariate normal distribution. If we let $u_n = y_t + z$ and $v_n = t - nm - s_k$, then using (8) we see that (7) can be bounded by

$$\leqq C \sum_{z \in A_t} \sum_{k=0}^{[(\beta \log t)/h]} P(\tau_1 > s_k; R_s = z)\, h\Psi(u, v).$$

We still have to sum over n. We have to estimate

$$\text{(9)} \qquad \sum_{n \in A_t} t^{d/2-\varepsilon} n^{-(d+1)/2} C \sum_{z} \sum_{k=0}^{[(\beta \log t)/h]} P(\tau_1 > s_k; R_s = z)\, h\Psi(u, v).$$

Using $\Psi(u,v)\leqq C$ and that $\frac{t}{n}$ is bounded by a constant on A_t (9) can be bounded by

$$\leqq C \sum_{n\in A_t} \left(\frac{t}{n}\right)^{d/2} t^{-\varepsilon} n^{-1/2} \sum_z \sum_{k=0}^{[\beta \log t/h]} h\cdot P(\tau_1 > s_k; R_s = z) \tag{10}$$

$$\leqq C \sum_{n\in A_t} t^{-\varepsilon} n^{-1/2} \sum_{k=0}^{[\beta \log t/h]} h P(\tau_1 > s_k)$$

where we also carried out the summation over z in the last step. Note that the distance between consecutive s_k's is h, hence we can approximate the sum over k by a Riemann integral and (10) can be bounded by

$$\leqq C \sum_{n\in A_t} t^{-\varepsilon} n^{-1/2} \int_0^\infty \mathrm{d}s P(\tau_1 > s)$$

$$\leqq C m t^{-\varepsilon} (t/m + t^{1/2+\varepsilon})^{-1/2}\, 2t^{1/2+\varepsilon}$$

where we used $m=\int_0^\infty \mathrm{d}s P(\tau_1>s)$. But the right-hand side is bounded by a constant $C\in(0,\infty)$. □

The next result provides us with an upper estimate on the probability that two different distinguished particles might hit for some t. Let $\Delta_{xy}(t)$ be the event that the triangles of two duals starting at x and y collide at time t. (For the definition of collision of triangles see the introduction.)

(11) Lemma. *For $d\geqq 3$ there is a constant $C<\infty$ so that*

$$P(\hat{\xi}_t^x(1)=\hat{\xi}_t^y(1) \text{ for some } t) \leqq P(\Delta_{xy}(t) \text{ for some } t) \leqq \frac{C}{\|x-y\|^{d-5/2}}.$$

Remark. A similar result is known for independent random walks (see e.g., Spitzer (1976)). Our estimate is slightly worse. We will replace the $d-2$ in the exponent in the independent case by $d-5/2$. This is due to the fact that our random walks are not quite independent.

Proof. The first inequality clearly holds. To prove the other inequality, set $y=0$ without loss of generality. Fix x, and let G_x and G_0 be two independent copies of the graphical gadget for the dual process. Use G_x to define $\hat{\xi}_t^x$ and G_0 to define $\hat{\xi}_t^0$ (see e.g., Griffeath (1979), p. 21 or the construction used at the beginning of Sect. 3b). We let the two dual processes evolve independently on G_x and G_0, respectively, as long as their triangles do not collide. After a collision of the triangles occurred, we use only G_x for both dual processes. Let $\hat{S}_t$ be the position of the embedded random walk at time t, and S_n its position after the nth jump. We break things down according to $t\leqq\|x\|^{2-1/d}$ and $t\geqq\|x\|^{2-1/d}$. To simplify notation we introduce $H=\Big\{$largest triangle by time $\|x\|^{2-1/d}$

$\left.\leqq \frac{2d}{\gamma} \log \|x\| \right\}$ where γ is the constant in (2.7). First, let $t \leqq \|x\|^{2-1/d}$. The idea behind the following estimate is that if the distinguished particles hit each other for some $t \leqq \|x\|^{2-1/d}$, then one of them must have left the ball with radius $\frac{1}{2}\left(\|x\| - \left|\frac{2d}{\gamma}\right| \log \|x\|\right)$ centered at the starting point of the corresponding dual process. Then

$$(12) \qquad P(\Delta_{x0}(t) \text{ for some } t) \leqq 2P\left(N(\|x\|^{2-1/d}) > \frac{2}{m} \|x\|^{2-1/d}\right)$$
$$+ 2P\left(\|\hat{S}_t\| \geqq \frac{1}{2}\left(\|x\| - \frac{2d}{\gamma} \log \|x\|\right); \right.$$
$$\left. N(\|x\|^{2-1/d}) \leqq \frac{2}{m} \|x\|^{2-1/d}; H\right)$$
$$+ \frac{4}{m} \|x\|^{2-1/d} P\left(\|X_1\| > \frac{2d}{\gamma} \log \|x\|\right).$$

$\|X_1\|$ denotes again the size of a triangle. For the first term on the right-hand side we use a large deviation estimate

$$(13) \qquad P\left(N(t) \geqq \frac{2}{m} t\right) \leqq C e^{-\gamma t}.$$

Hence

$$2P\left(N(\|x\|^{2-1/d}) \geqq \frac{2}{m} \|x\|^{2-1/d}\right) \leqq C \exp(-\gamma \|x\|^{2-1/d}).$$

For estimating the second term on the right-hand side of (12) note that by the reflection principle (applied to each coordinate separately)

$$(14) \qquad P(|\hat{S}_t^{(k)}| \geqq x \text{ for some } t \leqq t_0) \leqq 2P(|\hat{S}_{t_0}^{(k)}| \geqq x)$$

where $\hat{S}_t^{(k)}$ denotes the kth coordinate of $\hat{S}_t$.

If a random walk wants to leave a ball of radius r, at least one of its coordinates must be bigger than $\frac{r}{\sqrt{d}}$. This in mind we can estimate the second term on the right-hand side of (12) by

$$\leqq 2dP\left(|\hat{S}_t^{(1)}| \geqq \frac{1}{2\sqrt{d}}\left(\|x\| - \frac{2d}{\gamma} \log \|x\|\right); N(\|x\|^{2-1/d}) \leqq \frac{2}{m} \|x\|^{2-1/d}; H\right).$$

By (14) we can reduce the estimate to the case $n = \frac{2}{m} \|x\|^{2-1/d}$

$$\leqq 4dP\left(|S^{(1)}_{\frac{2}{m}\|x\|^{2-1/d}}| \geqq \frac{1}{2\sqrt{d}}\left(\|x\| - \frac{2d}{\gamma} \log \|x\|\right)\right).$$

We use Chebyshev's inequality to estimate this term

$$e^{\theta y} P(|S_n^{(1)}| > y) \leqq \mathbf{E} e^{\theta S_n^{(1)}} = [\Phi(\theta)]^n \tag{15}$$

where $\Phi(\theta)$ is the moment generating function of X_1. This gives for any $\theta > 0$

$$\leqq 4 d e^{-\theta \frac{1}{2\sqrt{d}}\left(\|x\| - \frac{2d}{\gamma}\log\|x\|\right)} e^{\frac{2}{m}\|x\|^{2-1/d} \log \Phi(\theta)}.$$

Taking $\theta = 2\|x\|^{\frac{1}{2d}-1}$ we can estimate by this

$$= 4d \exp\left\{-\left[\frac{\|x\|^{\frac{1}{2d}}}{\sqrt{d}} - \frac{2\sqrt{d}\,\|x\|^{\frac{1}{2d}}}{\gamma}\frac{\log\|x\|}{\|x\|} - \frac{2}{m}\|x\|^{2-\frac{1}{d}} C 4\|x\|^{\frac{1}{d}-2}\right]\right\}$$

where we used $\log \Phi(\theta) \leqq C\theta^2$ for θ small. This follows from the fact that the first derivative of $\Phi(\theta) = 0$ at $\theta = 0$ since the mean is, and the second derivative at 0 is finite since the variance is. Simplifying the exponent eventually gives

$$= 4d \exp\left\{-\|x\|^{1/2d}\left[\frac{1}{\sqrt{d}} - \frac{2\sqrt{d}}{\gamma}\frac{\log\|x\|}{\|x\|} - C\|x\|^{-1/2d}\right]\right\}.$$

If $\|x\|$ is large the last expression is

$$\leqq 4d \exp\left\{-\frac{1}{2\sqrt{d}}\|x\|^{1/2d}\right\}.$$

Finally, the third term in (12) can be estimated

$$\leqq C\frac{4}{m}\|x\|^{2-1/d}\frac{1}{\|x\|^{2d}} = C\frac{1}{\|x\|^{2d-2+1/d}}$$

by using the exponential estimates in Sect. 2. Putting things together the right-hand side of (12) is bounded by

$$\leqq C\exp\{-\gamma\|x\|^{2-1/d}\} + C\exp\left\{-\frac{1}{2\sqrt{d}}\|x\|^{1/2d}\right\} + C\frac{1}{\|x\|^{2d-2+1/d}}. \tag{16}$$

To get an upper bound on $P(\Delta_{x0}(t)$ for some $t \geqq \|x\|^{2-1/d})$ we will use that the dual processes are independent as long as their triangles do not collide, and hence we can use (4.1) to estimate the difference of the two duals. We will first control the duals at integer times and then show that there cannot get much wrong in between. More precisely, we will estimate (i) the probability of two duals (including their triangles) being further apart then $\frac{1}{2}n^{1/4d^2}$ at integer times n and (ii) the probability of the triangles overlapping in between.

To do (i) we need two estimates: The first one provides us with an upper bound on

$$(17)\qquad P(\|\xi_n^x(1)-\xi_n^y(1)\| \leqq n^{1/4d^2} \text{ for some } n \geqq \|x\|^{2-1/d}) \leqq \sum_{n \geqq \|x\|^{2-1/d}} P(\|\xi_n^x(1)-\xi_n^y(1)\| \leqq n^{1/4d^2})$$

We use (1) to estimate the summands. (The difference of $\xi_n^x(1)-\xi_n^y(1)$ obeys the Local Central Limit Theorem since they evolve independently as long as their triangles do not overlap. We will estimate the case where the triangles are too big afterwards.)

$$(18)\qquad \leqq \sum_{n \geqq \|x\|^{2-1/d}} C n^{-d/2+\varepsilon+1/4d} \leqq C \int_{\|x\|^{2-1/d}}^{\infty} s^{-d/2+\varepsilon+1/4d} \, ds = C(\|x\|^{2-1/d})^{-d/2+\varepsilon+1/4d+1} \leqq C\|x\|^{5/2-d}.$$

In the last step we chose $\varepsilon=1/4d$. As can be seen from the bound, this term gives the major contribution to the estimate. The second estimate shows that the triangles in either process cannot be too big.

$$(19)\qquad 2 \sum_{n \geqq \|x\|^{2-1/d}} P(\text{triangle at time } n > \tfrac{1}{4} n^{1/4d^2}) \leqq 2 \sum_{n \geqq \|x\|^{2-1/d}} P(\text{largest triangle up to time } n > \tfrac{1}{4} n^{1/4d^2})$$

The summand can be estimated according as whether $N(n)$, the number of jumps up to time n, is $>2n/m$ or $\leqq 2n/m$. Using (13) for the first case and the exponential estimate (2.7) for the second case, (19) is

$$\leqq 2 \sum_{n \geqq \|x\|^{2-1/d}} \left[C_2 e^{-\gamma_2 n} + \frac{2n}{m} C_1 e^{-\gamma_1 n^{1/4d^2}/4} \right]$$

C_1 and γ_1 are the constants in (2.7) and C_2, γ_2 the constants in (13). Summing over n, this can be bounded by

$$(20)\qquad \leqq C e^{-\gamma \|x\|^{\nu}}$$

for some constants $C, \gamma, \nu \in (0, \infty)$.

To see that not much can happen in between the integer times, observe that for the triangles to overlap in between, they have to grow fairly fast. At integer times n they are at least $n^{1/4d^2}/2$ units apart. So, if neither of their triangles grows more than $n^{1/4d^2}/4$ within one unit of time, they cannot collide. Since the boundary of the triangles grows like the boundary of a contact process, we estimate each coordinate separately by the right-edge process r_t of a contact process. There are d coordinates with two directions each, hence

(21)
$$\begin{aligned} &P(\text{triangles overlap within } [n, n+1) \text{ for some } n) \\ &\leqq \sum_{n \geqq \|x\|^{2-1/d}} 2d P(r_1 > n^{1/4d^2}/4) \\ &\leqq 2d \sum_{n \geqq \|x\|^{2-1/d}} C e^{-\delta n^{1/4d^2}} \leqq C e^{-\|x\|^{\nu}} \end{aligned}$$

for some constants C, γ, $\nu \in (0, \infty)$. Putting the pieces (18), (20) and (21) together proves the Lemma. □

5 Proof of Theorem 3

We will start with proving a result we call *convergence of trees*. Denote by $\eta \in (\mathbf{Z}^d)^n$ an ordered set of ancestors of size n shifted by the first ancestor so that $\hat{\xi}_t^x(1) = 0$, regardless of the starting point and where the distinguished particle actually landed. We go back k renewals and look at the *tree* growing out of this renewal point. We will call sometimes this point the *root*. Trees also have a certain *length*: If the root of the tree is located at (x, s) and t units of time have elapsed since we started the dual, then the length of the tree is $t - s$. We will show that we can find a countable partition of the sample space according to the length of the tree and where the distinguished particle lands so that the distribution of each atom has a limit. (This argument is similar to the one used in Durrett et al. 1989). Once this is shown it is easy to see that the distribution of the ancestor vector has a limit: each atom in the partition defines a certain ordered set of ancestors. Integrating over the length of the trees with a fixed ancestor vector finally proves the convergence of the distribution of the ordered set of ancestors.

Let $(\Omega_t, \mathscr{F}_t, P)$ be the probability space on which the graphical gadget is defined. Start the dual process at x and let it run until time t. The distinguished particle lands at $\hat{\xi}_t^x(1)$. Breaking things down according to the length of the tree growing out of its root and according to where the tree lands defines a countable partition of the sample space. More precisely: Denote by $l_t^{(1)}(\omega)$ the length of a tree with one renewal that was obtained by the above procedure on $\omega \in \Omega_t$. Denote by $[x]$ the integer part of x. Define the partition Π_t of Ω_t by considering two outcomes ω_1 and ω_2 to be in the same atom if and only if

(a) $[l_t^{(1)}(\omega_1)] = [l_t^{(1)}(\omega_2)]$ and

(b) $\eta_1 = \eta_2$.

This defines a countable partition of the sample space. The length of a tree has a density. It even has a limiting density as $t \to \infty$ by the renewal theorem. (The length of a tree is the amount of time since the last renewal, which has a limit (see e.g. Durrett 1990)). We will now show

(1) Lemma. *Let $\tau_{n,t}^{(k)}(\eta)$ be the atom of the trees whose length is in $(n, n+1]$ where $n+1 \leqq t$ with k renewals and ancestor vector η at time t. Then $\lim_{t \to \infty} P(\tau_{n,t}^{(k)}(\eta))$ exists.*

Proof. We will start with $k=1$. The measure of the set of trees whose length is in $(n, t]$ where $n=[t]$, and with a fixed number of renewals goes to zero as $t\to\infty$ by the exponential estimate (2.6). So we do not bother about this boundary effect. All trees in a certain atom $\tau_{n,t}^{(1)}(\eta)$ are of a fixed length $s\in(n, n+1]$. This means that the root of these trees are located at a site in $\mathbf{Z}^d\times[t-n-1, t-n)$. Given a certain length of the tree, the shape of the tree does no longer depend on where it is located because of spatial and temporal translation invariance of the graphical gadget, i.e.,

$$P(\tau_{n,t}^{(1)}(\eta)\,|\,l_t=s)=P(\tau_{n,t+h}^{(1)}(\eta)\,|\,l_{t+h}^{(1)}=s). \tag{2}$$

As we mentioned above, the length of a tree has a limiting density as $t\to\infty$ by the renewal theorem. Hence for t large we can find an $\varepsilon>0$ so that

$$\begin{aligned}
&|P(\tau_{n,t}^{(1)}(\eta))-P(\tau_{n,t+h}^{(1)}(\eta))|\\
&=\left|\int_n^{n+1} P(\tau_{n,t}^{(1)}(\eta)\,|\,l_t^{(1)}=s)\,P(l_t^{(1)}=s)\,\mathrm{d}s\right.\\
&\qquad\left.-\int_n^{n+1} P(\tau_{n,t+h}^{(1)}(\eta)\,|\,l_{t+h}^{(1)}=s)\,P(l_{t+h}^{(1)}=s)\,\mathrm{d}s\right|\\
&\leqq\int_n^{n+1} P(\tau_{n,t}^{(1)}(\eta)\,|\,l_t^{(1)}=s)\,|P(l_t^{(1)}=s)-P(l_{t+h}^{(1)}=s)|\,\mathrm{d}s\leqq\varepsilon
\end{aligned}$$

for all $h>0$ where we used (2) in the next to last step. By making t larger we can choose ε as small as desired by the convergence of the density of the length of the trees. This proves (1) for $k=1$.

To generalize this to the case of k renewals, define the atoms $\tau_{n,t}^{(k)}(\eta)$ now as containing those trees that have ancestor vector η, $l_t^{(k)}(\omega)\in(n, n+1]$, and exactly k renewals. The length of those trees also has a limiting density since the length of a tree is nothing else than a sum of interarrival times defined in the beginning of Sect. 3b and the amount of time since the last renewal. Both have a limit (see e.g. Durrett 1990). Replacing $^{(1)}$ by $^{(k)}$ in $\tau_{n,t}^{(1)}(\eta)$ above shows (1) for the general case. □

After establishing the convergence of the trees it is now straightforward to prove the convergence of the one-dimensional distribution by starting with any translation invariant measure μ. Let Γ_k be the event that at least one of the ancestors in the ancestor vector η lands on an occupied site. The subscript k refers to the number of renewals we look back. For $i=1$ or 2

$$\begin{aligned}
&\mathbf{E}\mathbf{E}^{\tau_{n,t}^{(k)}(\eta)}\mathbf{1}_{\{\mu(\eta)=i;\,\Gamma_k\}}\\
&\quad\leqq P(\xi_t^\mu(x)=i)=P(\text{first nonzero ancestor is } i)\\
&\quad\leqq\mathbf{E}\mathbf{E}^{\tau_{n,t}^{(k)}(\eta)}\mathbf{1}_{\{\mu(\eta)=i;\,\Gamma_k\}}+P(\Gamma_k^c)
\end{aligned}$$

where we used the shorthand notation $\{\mu(\eta)=i\}$ for the event that the first nonzero ancestor in the ancestor vector η is i under the measure μ. The convergence of the trees just proved can now be used to establish the convergence of $\mathbf{E}\mathbf{E}^{\tau_{n,t}^{(k)}(\eta)}\mathbf{1}_{\{\mu(\eta)=i;\,\Gamma_k\}}$. For a given ancestor vector η, the probability that $\{\mu(\eta)=i\}$ does not depend on t. The atoms have a limit distribution. So we average with respect to a limit distribution as $t\to\infty$. Since the number of particles in a tree is increasing in the number of renewals, it is clear that for every $\varepsilon>0$ we can find a $K>0$ so that $P(\Gamma_k^c)\leqq\varepsilon$ for all $k\geqq K$. Since ε is arbitrary, this proves the convergence.

For the convergence of the n-dimensional distribution we need again (4.1). From this and from the estimates on the triangles we can conclude that if we start with n particles, after a sufficiently long time the remaining particles (some of the n particles might coalesce) get separated so that they can be treated independently, and the convergence of the n-dimensional distributions follows from the one of the 1-dimensional distributions.

For proving this we first need a result that excludes the following situation: Two distinguished particles meet infinitely often without coalescing. We denote this event by A and prove

(3) Lemma. $P(A)=0$.

Proof. (3.5) and (3.6) show that (i) both duals are good infinitely often with probability one, and (ii) if the two distinguished particles are within a finite distance K, there is a positive probability that we can glue them together. It follows now from independence of what is going on before and after both duals were good, and the Borel-Cantelli Lemma that on the set where the two duals do not get separated, we can glue them together with probability 1. This proves that $P(A)=0$. □

The next result we need in proving the convergence of the n-dimensional distributions is the announced separation of the remaining distinguished particles for large times.

We will start two dual processes at x and y, respectively. We will prove in a series of Lemmas that on the set where they do not coalesce they get separated at least like $t^{1/10}$ for large t. We will do this by first showing that there is a positive probability that their distance is bigger than a fixed constant K for some t. Then we will show that if they start at a distance bigger than K, there is a positive probability that their distance is bigger than $t^{1/8}$ for all $t>0$. Using that they are both good infinitely often with probability 1 on the set where they do not coalesce, we can conclude with Borel-Cantelli that they get eventually separated. We start with

(4) Lemma. *For every $K>0$ we can find a $\delta_1>0$ so that*

$$P(\|\xi_t^x(1)-\xi_t^y(1)\|>K \text{ for some } t)\geqq\delta_1$$

independently of x and y.

Proof. We prove this by showing that the following event has positive probability. We require the particle located at x to survive for K units of time without giving birth. We enlarge the distance between x and y by giving birth in one coordinate direction. Then in at most K steps the distance is $\geqq K$ and both are good at the end. This has obviously positive probability. □

In the next step we are going to show

(5) Lemma. *There exists a large enough K so that the following is true: Let $\|x-y\|\geqq K$. On the set H where the triangles at time t are smaller than $C\log t$ we can find a $\delta_2>0$ so that*

$$P(\|\xi_t^x(1)-\xi_t^y(1)\|\geqq t^{1/8} \text{ for all } t\geqq 0; H)\geqq\delta_2.$$

Proof. We use the same idea as in the proof of (4.11). We split the proof into two parts. We start two independent copies at x and y and let them run until $K^{3/2}$. Since we consider everything on the set where the triangles are small, they will be independent until $t=K^{3/2}$. At this time we want the duals to be with high probability further away than $K-2K^{4/5}$. This will happen if both duals stay inside a ball of radius $K^{4/5}$ centered at their starting point. We will again use the reflection principle (4.14) and apply it to each coordinate separately. For this we denote by $\hat{S}_t$ the position of the embedded random walk at time t, and by S_n its position after the nth jump (we use the same notation as in the proof of (4.11)).

$$\begin{aligned}&P(\|\hat{\xi}_t^x(1)-\hat{\xi}_t^y(1)\| \leqq K-2K^{4/5} \text{ for some } 0 \leqq t \leqq K^{3/2}; H)\\ &\quad \leqq 2P(\|\hat{S}_t\| \geqq \tfrac{1}{2}K^{4/5} \text{ for some } 0 \leqq t \leqq K^{3/2}; H).\end{aligned}$$

Now we do the same calculation as in the proof of (4.11). We break things down according to $N(K^{3/2})>2K^{3/2}/m$ or $N(K^{3/2}) \leqq 2K^{3/2}/m$.

$$\begin{aligned}&\leqq 2P(\|\hat{S}_t\| \geqq K^{4/5}/2 \text{ for some } 0 \leqq t \leqq K^{3/2};\\ &\quad H; N(K^{3/2}) \leqq 2K^{3/2}/m)+2P(N(K^{3/2})>2K^{3/2}/m).\end{aligned}$$

For the first term we use the reflection principle (4.14) (applied to each coordinate separately). (4.13) takes care of the second term.

$$\leqq 4dP(|S^{(1)}_{2K^{3/2}/m}|>K^{4/5}/2\sqrt{d})+2dCe^{-\gamma K^{3/2}}.$$

(4.15) with $\theta=m/8C\sqrt{d}\,K^{0.7}$ yields to the estimate

$$\leqq Ce^{-\gamma K^{1/10}}.$$

We omit the details since the calculations are exactly the same as in the proof of (4.11). Hence with high probability their distance at times $\leqq K^{3/2}$ is bigger than $K-2K^{4/5}$.

To estimate $P(\|\hat{\xi}_t^x(1)-\hat{\xi}_t^y(1)\| \leqq t^{1/8}$ for some $t \geqq K^{3/2}; H)$ we do the same as in the proof of (4.11). We first estimate what happens at integer times and then show that not much can go wrong in between. The analogue of (4.17) and (4.18) is

$$\begin{aligned}&P(\|\hat{\xi}_n^x(1)-\hat{\xi}_n^y(1)\| \leqq n^{1/8} \text{ for some } n \geqq K^{3/2}; H)\\ &\quad \leqq \sum_{n \geqq K^{3/2}} P(\|\hat{\xi}_n^x(1)-\hat{\xi}_n^y(1)\| \leqq n^{1/8}; H)\\ &\quad \leqq \sum_{n \geqq K^{3/2}} Cn^{-d/2+\varepsilon+d/8} \leqq C\int_{K^{3/2}}^{\infty} s^{-d/2+\varepsilon+d/8}\,\mathrm{d}s \leqq CK^{-3/32}.\end{aligned}$$

We choose $\varepsilon=1/16$ in the last step. The estimate on the probability that the triangles overlap in between the integer times gives an additional term that is exponentially small in K. We omit the details. Hence

$$P(\|\xi_t^{\xi^x}(1)-\xi_t^{\xi^y}(1)\| \geqq t^{1/8} \text{ for all } t\geqq 0)$$
$$\geqq 1-CK^{-1/10}-2CK^{-3/32}.$$

By choosing K large enough we get the Lemma. □

Since both are good infinitely often with probability 1, they get eventually separated by the Borel-Cantelli Lemma on the set where they do not coalesce and where the triangles are not too big. We can conclude from the exponential estimate (2.7)

(6) $$P(\text{triangles} > C \log t \text{ for some } t > t_0)$$
$$\leqq A_1 \int_{t_0}^{\infty} t^{-A_2} \mathrm{d}t \leqq C t_0^{-2}$$

for arbitrary constants. Furthermore we proved that the distinguished particles hit only finitely many often. Let T be the last time they hit on the set where they do not coalesce, then

(7) $$P(T>t) \leqq \varepsilon_t$$

where $\varepsilon_t \to 0$ as $t \to \infty$.

The last two estimates together with Lemmas (4) and (5) prove that we have the following situation: They either coalesce early or they get separated at least like $t^{1/8}$. Since their triangles $\leqq C \log t$ with high probability, they will not collide for large t and we can treat them as independent. We will make use of this fact by proving that the n-dimensional distributions converge.

More precisely, we will define three "bad events" and show that the probability of their union goes to zero as t tends to infinity. Furthermore, on the complement, where the "good events" happen, everything will go well and the convergence will follow from the convergence of the 1-dimensional distributions as we will now explain. We start n dual processes. As we said the coalescing happens in the beginning. So, the first bad event will be $B_1=\{$some pair of the dual coalesce after $t_1\}$. We also need that the triangles behave well, so $B_2=\{$some triangle is bigger than $C\log t$ after $t_2\}$ is the second bad event. Finally, $B_3=\{$some pair of the dual comes closer than $t^{1/10}$ for some $t>t_3\}$ is the third bad event. We take $t^{1/10}$ in B_3 instead of $t^{1/8}$ as in (5) since we start counting time at 0 and we want those duals that meet finitely many times and get then separated, included into this set (note that $t^{1/10}$ and $(t-c)^{1/8}$ intersect for some $t>c$).

(5), (6), and (7) show that $P(B_1 \cup B_2 \cup B_3) \leqq \varepsilon$ where ε can be made as small as desired by choosing t_1 to t_3 sufficiently large. On the complement we are left with L particles where L is a random number. But those L dual processes are now independent and the convergence follows from the convergence of the 1-dimensional distributions. More precisely: Let $T=\max(t_1, t_2, t_3)$. Choose $T=\sqrt{t}$, then $P(B_1 \cup B_2 \cup B_3) \to 0$ as $t \to \infty$. On the complement $\bigcap_{k=1}^{3} B_k^c$, the duals that are left after $\sqrt{t}$ act independently. Their type is basically determined at $t-O(\log t)$ ($O(\log t)$ is the typical length of a tree with k renewals where k is fixed). Since $O(\log t) \ll t-\sqrt{t}$, the parts of the graphical gadget where the coalescing happens and where the type of a certain site is determined are far apart, so that they do not overlap. The convergence of each independently acting dual process follows from the convergence of the one dimensional distribution. This finally proves the weak convergence result, and we can turn to see what the limit distributions look like.

We now know that starting from any translation invariant measure the distribution of the system converges weakly to a limit. As we mentioned in the introduction the three and higher dimensional case contrasts the one and two dimensional case in the sense that we see coexistence of both types of particles in the limit. We will now characterize the limit distributions that arise if we start the system with a translation invariant and ergodic measure μ. We will show that

$$\text{(8)} \qquad \xi_t^\mu \Rightarrow \nu_\theta \quad \text{for some } \theta \text{ as } t \to \infty$$

where $\{\nu_\theta\}_{0 \leqq \theta \leqq 1}$ forms a one-parameter family of invariant measures that are translation invariant and ergodic.

We just proved the convergence of μ. That the limit is translation invariant is also clear from the choice of μ and since the dynamic is translation invariant. θ characterizes the density of 1's in the limit:

$$\lim_{t \to \infty} P(\xi_t^\mu(0)=1) = \lim_{t \to \infty} P(\xi^\mu(\hat{\xi}_t^0)=1)$$
$$= \lim_{t \to \infty} \sum_{x \in \mathbf{Z}^d} P(\xi(\hat{\xi}_t^0 - x)=1)\, P(\hat{\xi}_t^0(1)=x)$$

where $\{\xi^\mu(\hat{\xi}_t^0)=1\}$ is a short hand notation for the event that the first nonzero entry in the ancestor vector $\hat{\xi}_t^0$ lands on a 1. From the convergence of the trees and the continuous mapping theorem we can conclude that there is a $0 \leqq \theta \leqq 1$ so that $P(\xi(\hat{\xi}_t^0 - x)=1) \to \theta$. Hence by translation invariance

$$\lim_{t \to \infty} P(\xi_t^\mu(0)=1) = \theta \lim_{t \to \infty} P(\hat{\xi}_t^0(1) \text{ lands somewhere}) = \theta P(\Omega_\infty).$$

What is left to show is that $\{\nu_\theta\}_{0 \leqq \theta \leqq 1}$ are mutually singular and ergodic. From this it follows that

$$\text{(9)} \qquad (\mathscr{I} \cap \mathscr{S})_e = \{\nu_\theta : 0 \leqq \theta \leqq 1\}$$

where $\mathscr{I}$ denotes the set of invariant measures and $\mathscr{S}$ the set of translation invariant measures since the translation invariant and ergodic measures are the extreme points of the set of translation invariant measures (Dynkin 1978). Following Sect. 11 b in Durrett (1988) we will first prove that

(10) Lemma. $\{\nu_\theta\}_{0 \leqq \theta \leqq 1}$ *are mutually singular.*

Proof. Let $\xi \equiv \xi_0 = \nu_\theta$ for some $0 \leqq \theta \leqq 1$. Call $\zeta(x) = \mathbf{1}_{\{\xi(x)=1\}} - P(\xi(x)=1)$, and $S_t = \sum_{x \in \Lambda_t} \zeta(x)$ where $\Lambda_t = [-t, t]^d$. We will first show by applying Chebyshev's

Inequality that $\frac{|S_t|}{(2t+1)^d} \to 0$ in probability. For this we need an estimate on the second moment of S_t. The ergodic theorem will then show that the convergence actually occurs almost surely, which will prove the Lemma. We begin with an estimate on the second moment of S_t:

$$\mathbf{E}S_t^2 = \mathbf{E}\Big(\sum_{x\in\Lambda_t} \zeta(x)\Big)^2$$
$$= \sum_{x\in\Lambda_t} \mathbf{E}(\zeta(x))^2 + \sum_{\substack{x,y\in\Lambda_t\\ x\neq y}} \mathbf{E}(\zeta(x)\,\zeta(y))$$

can be estimated by using

$$\mathbf{E}(\zeta(x))^2 = P(\xi(x)=1) - [P(\xi(x)=1)]^2 \leqq 1.$$

With $\Delta_{xy}(t)$ being the event that the triangles of two duals starting at x and y collide at time t, and with (4.11)

$$\mathbf{E}(\zeta(x)\,\zeta(y)) \leqq P(\Delta_{xy} \text{ for some } t) \leqq \frac{C}{\|x-y\|^{d-5/2}},$$

and

$$\sum_{\substack{x,y\in\Lambda_t\\ x\neq y}} \frac{C}{\|x-y\|^{d-5/2}} \leqq (2t+1)^d \sum_{m=1}^{2t+1} 2d(2m+1)^{d-1} \frac{1}{m^{d-5/2}}$$
$$= 2dC(2t+1)^d \sum_{m=1}^{2t+1} \left(\frac{2m+1}{m}\right)^d \frac{m^{5/2}}{2m+1} \leqq 2dC(2t+1)^d\, 3^d \sum_{m=1}^{2t+1} \frac{m^{3/2}}{2}$$
$$\leqq 2dC(2t+1)^d\, 3^d \tfrac{1}{2}(2t+1)^{3/2}(2t+1) \leqq C(2t+1)^{d+5/2}$$

Using Chebyshev's Inequality and the estimates on the second moments now gives

$$P(|S_t| \geqq \varepsilon(2t+1)^d) \leqq \frac{\mathbf{E}S_t^2}{\varepsilon^2(2t+1)^{2d}}$$
$$\leqq \frac{1}{\varepsilon^2(2t+1)^{2d}} \{(2t+1)^d + C(2t+1)^{d+5/2}\}$$
$$\leqq \frac{C}{(2t+1)^{d-5/2}} \to 0$$

as $t\to\infty$ if $d \geqq 3$. This tell us that

$$\text{(11)} \qquad \frac{|S_t|}{(2t+1)^d} \to 0 \quad \text{in probability.}$$

Furthermore, the multiparameter ergodic theorem (see e.g., Dunford and Schwartz 1958) says that

$$(12)\qquad \frac{1}{(2t+1)^d}\sum_{x\in\Lambda_t}\mathbf{1}_{\{\xi_t(x)=1\}}\to \mathbf{E}(\mathbf{1}_{\{\xi(0)=1\}}|\mathscr{A})\quad \nu_\theta\text{-a.s.}$$

where $\mathscr{A}$ denotes the invariant σ-algebra. Comparing (11) with the conclusion of the ergodic theorem shows that

$$\frac{1}{(2t+1)^d}\sum_{x\in\Lambda_t}\mathbf{1}_{\{\xi(x)=1\}}\to\theta\quad \nu_\theta\text{-a.s.},$$

so the ν_θ are mutually singular. □

By this method we can prove more:

(13) Lemma. $\{\nu_\theta\}_{0\leqq\theta\leqq1}$ *are ergodic.*

Proof. Let $B(F_1,F_2)=\{\xi(x)=1$ for all $x\in F_1;\ \xi(x)=2$ for all $x\in F_2\}$ where F_1 and F_2 are finite subsets in $\mathbf{Z}^d$. As we mentioned in the beginning of Sect. 4, these sets form a convergence determining class. Let $F=F_1\cup F_2$. As before we will first show that

$$(14)\qquad \frac{1}{(2t+1)^d}\sum_{x\in\Lambda_t}\mathbf{1}_{B(F+x)}\to P(B(F))$$

in probability, where $B(F+x)=B(F_1+x,F_2+x)$ and Λ_t is the set defined above. As before, let $\zeta(F)=\mathbf{1}_{B(F)}-P(B(F))$ and $S_t=\sum_{x\in\Lambda_t}\zeta(F+x)$. Following the proof where F consists only of two points, we need an estimate on $\mathbf{E}[\zeta(F+x)\,\zeta(F+y)]$ where we will first assume that $(x+F)\cap(y+F)=\emptyset$. Then we can bound

$$\begin{aligned}&\mathbf{E}[\zeta(F+x)\,\zeta(F+y)]\\ &\leqq P(\Delta_{uv}(t)\colon \text{for some } u\in x+F,\ v\in y+F, \text{ and for some } t>0)\\ &\leqq\binom{|F|}{2}\sup\{P(\Delta_{uv}(t)\text{ for some } t>0)\colon (u,v)\in(x+F,y+F)\}\end{aligned}$$

where $|F|$ denotes the cardinality of F. Using (4.11), the last quantity is

$$\leqq\binom{|F|}{2}\sup\left\{\frac{C}{\|u-v\|^{d-5/2}}\colon (u,v)\in(x+F,y+F)\right\}$$

Then

$$\begin{aligned}P(|S_t|\geqq\varepsilon(2t+1)^d)&\leqq\frac{\mathbf{E}|S_t|^2}{\varepsilon^2(2t+1)^{2d}}\\ &\leqq C\binom{|F|}{2}\frac{1}{(2t+1)^{d-5/2}}\to 0\end{aligned}$$

as $t\to\infty$ if $d\geqq3$. If $(x+F)\cap(y+F)\neq\emptyset$ then we can use that $\mathbf{E}[\zeta(x+F)\,\zeta(y+F)]\leqq1$. For a fixed x there are a finite number of y's so that

$(x+F)\cap(y+F)\neq\emptyset$. Let this number be K. K does not depend on t or x. The number of terms we get from this contribution in the summation is therefore of order $O(\Lambda_t)$ and thus this term as well as the term coming from the estimate of $\mathbf{E}[\zeta(x+F)^2]$ vanishes in the Chebyshev estimate. The multiparameter ergodic theorem shows

$$\left(\frac{1}{2t+1}\right)^d \sum_{x\in\Lambda_t} \mathbf{1}_{B(F+x)} \to \mathbf{E}(\mathbf{1}_{B(F)}|\mathscr{A}) \quad \nu_\theta\text{-a.s.} \tag{15}$$

Hence by the same argument as above

$$\mathbf{E}(\mathbf{1}_{B(F)}|\mathscr{A})=P(B(F)) \quad \nu_\theta\text{-a.s.}$$

for finite $F=(F_1,F_2)$. We will now show that $\mathscr{A}$, the invariant σ-algebra is trivial from which the Lemma follows. Using the monotone class theorem it suffices to take a sequence of finite dimensional sets A_n where the A_n are increasing and $A=\bigcup_{n=1}^{\infty} A_n$ and where each A_n is of the form $B(F_1^{(n)}, F_2^{(n)})$, and $F_1^{(n)}$ and $F_2^{(n)}$ are both increasing. The monotone convergence theorem gives

$$\lim_{N\to\infty} P\left(B\left(\bigcup_{n=1}^{N} A_n\right)\right) = P\left(B\left(\bigcup_{n=1}^{\infty} A_n\right)\right)=P(B(A)).$$

Using this we can conclude

$$\begin{aligned}\mathbf{E}(\mathbf{1}_{B(A)}|\mathscr{A}) &= \lim_{N\to\infty} \mathbf{E}\left(\mathbf{1}_{B\left(\bigcup_{n=1}^{N} A_n\right)}\middle|\mathscr{A}\right)\\ &= \lim_{N\to\infty} P\left(B\left(\bigcup_{n=1}^{N} A_n\right)\right)=P(B(A))\end{aligned}$$

where the middle equality follows from (15). Since $P(B(A))$ is not a random variable, $\mathscr{A}$ is trivial, and $\{\nu_\theta\}_{0\leq\theta\leq1}$ are therefore ergodic. □

Acknowledgement. This work is taken from my doctoral dissertation at Cornell University under supervision of Rick Durrett. I am grateful to him for his advice and encouragement during this time. I also thank the referees for careful reading and pointing out inaccuracies in the first version.

References

Bezuidenhout, C., Grimmett, G.: The critical contact process dies out. Ann. Probab. **18**, 1462–1482 (1990)

Billingsley, P.: Probability and measure, 2nd edn. New York: Wiley 1986

Chung, K.L.: A course in probability theory, 2nd edn. New York: Academic Press 1974

Dunford, N., Schwartz, J.T.: Linear operators, Part I: general theory. New York: Interscience 1958

Durrett, R.: Lecture Notes on Particle Systems and Percolation. Californien: Wadsworth & Brooks 1988

Durrett, R.: Probability: Theory and examples. Californien: Wadsworth & Brooks 1990

Durrett, R., Schonmann, R., Tanaka, N.: The contact process on a finite set, III: The critical case. Ann. Probab. **17**, 1303–1321 (1989)

Dynkin, E.B.: Sufficient statistics and extreme points. Ann. Probab. **6**, 705–730 (1978)

Griffeath, D.: Additive and cancellative interacting particle systems. (Lect. Notes Math. vol. 724) Berlin Heidelberg New York: Springer 1979

Harris, T.E.: Nearest neighbor Markov interaction processes on multidimensional lattices. Adv. Math. **9**, 66–89 (1972)

Harris, T.E.: On a class of set valued Markov processes. Ann. Probab. **4**, 175–194 (1976)

Holley, R., Liggett, T.: Ergodic theorems for weakly interacting particle systems and the voter model. Ann. Probab. **3**, 643–663 (1975)

Kuczek, T.: The central limit theorem for the right edge of supercritical oriented percolation. Ann. Probab. **17**, 1322–1332 (1989)

Liggett, T.M.: Interacting particle systems. Berlin Heidelberg New York: Springer 1985

Spitzer, F.: Principles of random walk, 2nd edn. Berlin Heidelberg New York: Springer 1976

Stone, C.: On local and ratio limit theorems. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. II, part 2. Berkeley and Los Angeles, University of California Press, pp. 217–224, 1967