The Annals of Probability
1989, Vol. 17, No. 4, 1333–1366

COALESCING RANDOM WALKS AND VOTER MODEL CONSENSUS TIMES ON THE TORUS IN $\mathbb{Z}^d$

BY J. T. COX[1]

Syracuse University

Let η_t be the basic voter model on $\mathbb{Z}^d$ and let $\eta_t^{(N)}$ be the voter model on $\Lambda(N)$, the torus of side N in $\mathbb{Z}^d$. Unlike η_t, $\eta_t^{(N)}$ (for fixed N) gets trapped with probability 1 as $t \to \infty$ at all 0's or all 1's. We examine the asymptotic growth of these trapping or *consensus* times $\tau^{(N)}$ as $N \to \infty$. To do this we obtain limit theorems for coalescing random walk systems on the torus $\Lambda(N)$, including a new hitting time limit theorem for (noncoalescing) random walk on the torus.

1. Introduction. Infinite particle systems are stochastic processes which model the behavior of large systems of stochastically interacting components. Typically the components are located at the points (sites) of a set $\Lambda \subset \mathbb{Z}^d$ and can be in several different states, the simplest case being that of two states, say 0 and 1. The state of the system at time t is η_t, an element of $\{0,1\}^\Lambda$; $\eta_t(x)$ is the state of the component at site x at time t.

From the applied point of view, one is interested in the behavior of such processes when Λ is finite but very large. The usual approach is to replace Λ with $\mathbb{Z}^d$ and then study the infinite system. This leads to a rich and beautiful theory. Moreover, it is generally believed that infinite systems provide good approximations for large finite systems [see Dobrushin (1971) for a discussion of this point]. We are interested here in trying to better understand what this notion of approximation means.

To fix the ideas a little more clearly, let η_t be an interacting particle system on $\mathbb{Z}^d$ and let η_t^Λ be some suitable version of η_t restricted to a finite set $\Lambda \subset \mathbb{Z}^d$. If t is fixed and $\Lambda \uparrow \mathbb{Z}^d$, then it is usually the case that η_t^Λ converges weakly to η_t. On the other hand, if we fix Λ and let $t \to \infty$, then things are different. For instance, the contact process of Harris (1974) has a single trap, the element which is identically zero. The finite contact process η_t^Λ will almost surely hit this trap, no matter what the initial configuration. This is not true for η_t; if the infection parameter is sufficiently large, η_t has a nondegenerate equilibrium. Thus the ergodic behavior of η_t differs from that of η_t^Λ in a fundamental way. The problem is that simply letting $t \to \infty$ is too crude. One still expects that if Λ is large and t is large *but not too large*, then locally the finite and infinite systems should look the same. Consequently it is of interest to study η_t^Λ as both $t \to \infty$ and $\Lambda \uparrow \mathbb{Z}^d$.

Another point of view is that the infinite systems are the primary objects of study and that finite systems are approximations to them. This is especially

Received September 1987; revised August 1988.
[1]Research supported in part by NSF Grant DMS-86-01713.
AMS 1980 *subject classification.* Primary 60K35.
Key words and phrases. Infinite particle systems, finite particle systems, random walks, voter model.

relevant if one hopes to understand the behavior of infinite systems through computer simulations. [Durrett (1987) is a good source of simulations of many interacting particle systems.] Of course what is simulated is a *finite* system, so we are back to the basic difficulty: How long can you watch a finite system before it *knows* it is a finite system?

A natural first step in looking at this problem is to study the asymptotics of trapping times which clearly distinguish the finite and infinite systems. This has been done for the contact process. The first work along these lines appears in Griffeath (1981), while more recent work is in Cassandro, Galves, Olivieri and Vares (1984), Schonmann (1985), Durrett and Liu (1988) and Durrett and Schonmann (1988).

In this paper we will focus on trapping times of the voter model (and its dual, coalescing random walk) of Clifford and Sudbury (1973) and Holley and Liggett (1975). Voter model trapping times on various finite graphs have been studied by Donnelly and Welsh (1983), but they dealt only briefly with the torus in $\mathbb{Z}^d$ which is our main interest here. We start by defining our process and giving a little background information about it.

If $\Lambda \subset \mathbb{Z}^d$ and $p^\Lambda(x, y)$ is a stochastic matrix on Λ, then the corresponding voter model is η_t^Λ, the Markov process with state space $\{0,1\}^\Lambda$ which makes transitions

$$\eta_t^\Lambda(x) \to 1 - \eta_t^\Lambda(x) \quad \text{at rate} \quad \sum_{y \in \Lambda} p^\Lambda(x, y) 1\big(\{\eta_t^\Lambda(x) \neq \eta_t^\Lambda(y)\}\big).$$

[$1_A = 1(A)$ is the indicator function of A.] That is, each voter waits a random time which is exponential with parameter 1, selects another voter according to p^Λ and adopts the opinion of that voter. Observe that there are two traps, all 0's and all 1's. For $\Lambda = \mathbb{Z}^d$ we drop the superscript Λ.

The infinite system. Let $\Lambda = \mathbb{Z}^d$ and let $p(x, y)$ be the transition function of simple symmetric random walk on $\mathbb{Z}^d$, $p(x, y) = (2d)^{-1} 1\{|x - y| = 1\}$. Write η_t^μ if the initial distribution is μ and let $\mathscr{L}(\eta_t^\mu)$ be the law of η_t^μ. To describe the basic ergodic theory of η_t, due to Clifford and Sudbury (1973) and Holley and Liggett (1975), let $\mathscr{I}$ be the set of invariant probability measures for η_t and let $\mathscr{I}_e$ denote the set of extreme points of $\mathscr{I}$. For $0 \le \theta \le 1$ let μ_θ be product measure with density θ, i.e., $\mu_\theta\{\eta(x) = 1\} = \theta$ for all $x \in \mathbb{Z}^d$, and let $\Rightarrow$ denote weak convergence of probability measures.

THEOREM 0. *If $d \le 2$, then $\mathscr{I}_e = \{\mu_0, \mu_1\}$ and $\mathscr{L}(\eta_t^{\mu_\theta}) \Rightarrow (1-\theta)\mu_\theta + \theta\mu_1$ as $t \to \infty$. If $d \ge 3$, then there are probability measures ν_θ, $0 \le \theta \le 1$, such that $\mathscr{I}_e = \{\nu_\theta, 0 \le \theta \le 1\}$ and $\mathscr{L}(\eta_t^{\mu_\theta}) \Rightarrow \nu_\theta$ as $t \to \infty$.*

This is only a special case of Theorem V.1.8 in Liggett (1985), and much more is known. For instance, the domain of attraction of each ν_θ can be described. In particular, if μ is a translation invariant, shift ergodic measure with $\int \eta(x)\, d\mu(\eta) = \theta$, then $\mathscr{L}(\eta_t^\mu) \Rightarrow \nu_\theta$. [See Liggett (1985), Theorem V.1.9 for a complete discussion and details.]

Theorem 0 indicates that *clustering* occurs for $d \le 2$, *stability* for $d \ge 3$. The clustering is studied in Bramson and Griffeath (1980a) and Arratia (1979) ($d = 1$)

and Cox and Griffeath (1988) ($d = 2$). The macroscopic structure of the invariant measures ν_θ ($d \geq 3$) is considered in Bramson and Griffeath (1979). In all dimensions it is the case that, with probability 1, if $0 < \theta < 1$, then $\eta_t^{\mu_\theta}(0)$ changes state infinitely often as $t \to \infty$, so even in one and two dimensions the infinite voter model avoids being trapped.

The finite systems. We will consider a sequence of finite systems by taking $\Lambda(N) = \mathbb{Z}^d \cap [-N/2, N/2)^d$, $N = 2, 4, \ldots$. We will regard $\Lambda(N)$ as a torus and write $p^{(N)}(x, y)$ for the transition function of simple symmetric random walk on $\Lambda(N)$. That is, if x, y are in $\Lambda(N)$, then

$$p^{(N)}(x, y) = \sum_z p(x, z) 1(y \equiv z \bmod(N)).$$

We will write $\eta_t^{(N)}$ for $\eta_t^{\Lambda(N)}$. Since $\eta_t^{(N)}$ is a continuous time Markov chain on a finite state space, it is easy to see that no matter what the initial state or dimension, $\eta_t^{(N)}$ gets trapped at all 0's or all 1's with probability 1. This contrasts sharply with the behavior of Theorem 0.

ASSUMPTION. From now on, unless otherwise indicated, η_t will have initial distribution μ_θ and $\eta_t^{(N)}$ will have initial distribution μ_θ restricted to $\Lambda(N)$.

We are interested in determining the asymptotic growth of the "consensus times"

$$\tau^{(N)} = \inf\{t \geq 0 \colon \eta_t^{(N)} \equiv 0 \text{ or } 1 \text{ on } \Lambda(N)\}.$$

To describe our results we require a little more notation. Let $p_n^{(N)}(x, y)$ be the nth iterate of $p^{(N)}(x, y)$, and define

$$s_N = \begin{cases} N^2, & d = 1, \\ N^2 \log N, & d = 2, \\ N^d, & d \geq 3, \end{cases} \tag{1.1}$$

$$G = \begin{cases} \frac{1}{6}, & d = 1, \\ 2/\pi, & d = 2, \\ \sum_{n=0}^{\infty} p_n(x, y), & d \geq 3, \end{cases} \tag{1.2}$$

$$q_{n,k}(t) = \sum_{j=k}^{n} \frac{(-1)^{j+k}(2j-1)(j+k-2)!\binom{n}{j}}{k!(k-1)!(j-k)!\binom{n+j-1}{j}} \exp\left(-t\binom{j}{2}\right), \tag{1.3}$$

$$q_{\infty,k}(t) = \sum_{j=k}^{\infty} \frac{(-1)^{j+k}(2j-1)(j+k-2)!}{k!(k-1)!(j-k)!} \exp\left(-t\binom{j}{2}\right), \tag{1.4}$$

for $t > 0$, $1 \leq k \leq n < \infty$. The $q_{n,k}$ have a simple probabilistic representation that greatly facilitates the derivation of elementary properties of the $q_{n,k}$.

Namely, if we let $D_t = D(t)$ be a Markov chain on $\{1, 2, \dots\}$ with transition mechanism

$$n \to n-1 \quad \text{at rate} \binom{n}{2},$$

then $P_n[D_t = k] = q_{n,k}(t)$. See Tavaré (1984) or Cox and Griffeath (1988) for more on this.

THEOREM 1. *There are random variables τ depending on the dimension d such that as $N \to \infty$,*

$$\tau^{(N)}/s_N \Rightarrow \tau \quad \text{and} \quad E\left[\tau^{(N)}/s_N\right] \to E[\tau]. \tag{1.5}$$

If $d \geq 2$, then

$$P[\tau \leq s] = \sum_{k=1}^{\infty} \left[\theta^k + (1-\theta)^k\right] q_{\infty,k}(2s/G), \qquad s \geq 0, \tag{1.6}$$

and $E[\tau] = -G[\theta \log \theta + (1-\theta)\log(1-\theta)]$.

In contrast to Theorem 0, Theorem 1 singles out $d = 1$ and lumps $d = 2$ with $d \geq 3$. Furthermore, it shows that s_N determines an important time scale for $\eta_t^{(N)}$. For further evidence of this consider the "density" process

$$\Delta_t^{(N)} = N^{-d} \sum_{x \in \Lambda(N)} \eta_t^{(N)}(x)$$

and let Y_t be the one-dimensional diffusion on $[0, 1]$ with initial point θ and generator

$$\frac{1}{2}\gamma(1-\gamma)\frac{d^2}{d\gamma^2}$$

(both 0 and 1 are accessible traps). We will refer to Y_t as the Wright–Fisher diffusion. The following result shows that the particle density on $\Lambda(N)$ fluctuates like the Wright–Fisher diffusion with time scale s_N.

THEOREM 2. *If $d \geq 2$, then as $N \to \infty$,*

$$\Delta_{ts_N}^{(N)} \Rightarrow Y_{2t/G} \tag{1.7}$$

as processes.

The appearance of the Wright–Fisher diffusion transition function controlling the density of $\eta_t^{(N)}$ is suggestive; consider the following argument. At time ts_N, since $\eta_{ts_N}^{(N)}$ has density $\Delta_{ts_N} \approx Y_{2t/G}$, the distribution of $\eta_{ts_N}^{(N)}$ should be approximately $\nu_{\theta'}$ with $\theta' = Y_{2t/G}$. This is in fact correct: If $d \geq 3$ and $t_N/s_N \to t \in [0, \infty]$ as $N \to \infty$, then

$$\mathscr{L}\left(\eta_{t_N}^{(N)}\right) \Rightarrow \int_{[0,1]} P\left[Y_{2t/G} \in d\theta'\right] \nu_{\theta'}. \tag{1.8}$$

The precise result is

THEOREM 3. *Assume $d \geq 3$, $A \subset \mathbb{Z}^d$ is finite and ζ is fixed. If $t_N \to \infty$ and $t_N/N^d \to t \in [0, \infty]$ as $N \to \infty$, then*

$$P\left[\eta_{t_N}^{(N)}(x) = \zeta(x), x \in A\right] \to \int_{[0,1]} P\left[Y_{2t/G} \in d\theta'\right] \nu_{\theta'}[\eta(x) = \zeta(x), x \in A].$$

This topic is carried further in Cox and Greven (1988).

The proofs of Theorems 1–3 are based on some new theorems for coalescing random walk on the torus, which we feel are of interest in their own right. The coalescing random walk system ξ_t is easily defined. Its state space is the set of all subsets of $\mathbb{Z}^d$ and $\xi_t(A)$ is the set of occupied sites at time t when the initial state is $A \subset \mathbb{Z}^d$; we write $\xi_t(x)$ for $\xi_t(\{x\})$. Each particle independently executes simple symmetric rate 1 continuous time random walk on $\mathbb{Z}^d$, except that when a particle lands on a site already occupied by a particle the two particles coalesce into one. In the obvious way we can define $\xi_t^{(N)}$, coalescing random walk on the torus $\Lambda(N)$. Unless otherwise noted we assume that ξ_t is constructed using the graphical representation of Harris (1978) [see also Griffeath (1979)]. This means, in particular, that all $\xi_t^{(N)}(B)$, $B \subset \mathbb{Z}^d$, are defined on a common probability space with

$$\xi_t^{(N)}(B_1 \cup B_2) = \xi_t^{(N)}(B_1) \cup \xi_t^{(N)}(B_1).$$

There is a duality relation (see Section 4) between η_t and ξ_t ($\eta_t^{(N)}$ and $\xi_t^{(N)}$) that transforms certain questions about the voter model into questions concerning the *cardinality* of the coalescing random walk system. As might be expected, the behavior of the finite system $\xi_t^{(N)}$ differs from that of infinite system ξ_t. For example, suppose we start $\xi_t^{(N)}$ with two particles in $\Lambda(N)$. Then in any dimension the two walks are eventually bound to meet, i.e., eventually $\xi_t^{(N)}$ has cardinality 1; of course this is not the case for the infinite system if $d \geq 3$. Analysis of the voter model on the torus leads us naturally to the question of how long it takes random walks on the torus to collide.

Random walk on the torus, and various related models, have been studied for some time. Montroll (1969) and den Hollander and Kasteleyn (1982) are good sources for references to this literature. As far as we know Theorem 4 below is new. To state it let $X_t^{(N)}$, $t \geq 0$, be simple symmetric rate 1 continuous time random walk on the torus $\Lambda(N)$ and let $H^{(N)}$ be the hitting time of the origin, $H^{(N)} = \inf\{t \geq 0\colon X_t^{(N)} = 0\}$. Our result is

THEOREM 4. *Assume $d \geq 2$, $a_N \to \infty$ and $a_N = o(N)$ as $N \to \infty$. For $d = 2$ assume in addition that $a_N\sqrt{\log N}/N \to \infty$. Then, uniformly in $t \geq 0$ and $x \in \Lambda(N)$ with $|x| \geq a_N$,*

$$P_x\left[H^{(N)}/s_N > t\right] \to \exp[-t/G].$$

Note that if $X_0^{(N)}$ is uniformly distributed over $\Lambda(N)$, then Theorem 4 implies that

$$P\left[H^{(N)}/s_N > t\right] \to \exp[-t/G],$$

a result of Flatto, Odlyzko and Wales (1985) (see Theorems 6.1 and 6.2). The case $d = 1$ is different; it is easy to guess (and prove) what happens in this case [again, see Flatto, Odlyzko and Wales (1985)].

We can view Theorem 4 as a result concerning coalescing random walk on the torus. For $x_1, x_2 \in \Lambda(N)$ let $x_1 + x_2$ denote addition on $\Lambda(N)$. Then we may regard $\xi_t^{(N)}(x_1) - \xi_t^{(N)}(x_2)$ as a rate 2 random walk on $\Lambda(N)$ up until the time that the random walks meet. With $|A|$ = cardinality of A, Theorem 4 implies

$$(1.9)\qquad P\Big[\big|\xi_{ts_N}^{(N)}(\{x_1, x_2\})\big| = 1\Big] = P_{x_1 - x_2}\big[H^{(N)}/s_N > 2t\big] \to 1 - e^{-2t/G}$$

as $N \to \infty$, provided that $|x_1 - x_2| \geq a_N$.

The keys to proving Theorems 1 and 2 are extensions of (1.9) which handle coalescing random walk systems starting with *more* than two walks.

THEOREM 5. *Assume $d \geq 2$, $T > 0$ and $n \geq 2$, and a_N satisfies the assumptions of Theorem 4. Then, uniformly in $0 \leq t \leq T$ and $A^{(N)} = \{x_1, x_2, \dots, x_n\} \subset \Lambda(N)$ with $|x_i - x_j| \geq a_N$ for $i \neq j$,*

$$(1.10)\qquad P\Big[\big|\xi_{ts_N}^{(N)}(A^{(N)})\big| = k\Big] \to q_{nk}(2t/G), \qquad 1 \leq k \leq n.$$

With this result and a "patch" similar to the proposition in Bramson, Cox and Griffeath (1986), page 615 we can "fill up" the torus with walks and obtain asymptotics for the time it takes $\xi_t^{(N)}(\Lambda(N))$ to coalesce to j walks.

THEOREM 6. *Let $\xi_0^{(N)} = \Lambda(N)$ and let $\sigma_j^{(N)} = \inf\{t \geq 0\colon |\xi_t^{(N)}| = j\}$. There are random variables σ_j such that for $j = 1, 2, \dots$ as $N \to \infty$,*

$$(1.11)\qquad \sigma_j^{(N)}/s_N \Rightarrow \sigma_j \quad \text{and} \quad E\big[\sigma_j^{(N)}/s_N\big] \to E[\sigma_j].$$

If $d = 1$, then

$$(1.12)\qquad E\big[e^{-\alpha\sigma_1}\big] = \frac{\sqrt{\alpha}}{\sinh\sqrt{\alpha}}, \qquad \alpha > 0.$$

If $d \geq 2$, then

$$(1.13)\qquad P\big[\sigma_j \leq s\big] = \sum_{k=1}^{j} q_{\infty,k}(2s/G), \qquad s \geq 0.$$

In all dimensions $E[\sigma_1] = G$.

One of the themes of this work is that the spatial dependence in our models "washes out" in an appropriate limit. We point out that Kingman's coalescent [Kingman (1982)], a process *without* spatial dependence, is lurking in the background [see Cox and Griffeath (1988) where this matter is more fully explored].

Having described the main results we now state how the rest of the paper is organized. In Section 2 we derive the main probability estimates for random walk on the torus and prove Theorem 4. In Section 3 we prove some probability limit theorems for coalescing random walk on the torus, including Theorem 5. We follow the methods of Cox and Griffeath (1988) very closely in this section. In Section 4 we estimate the expected number of random walks left in $\xi_{ts_N}^{(N)}(\Lambda(N))$, using techniques of Bramson and Griffeath (1980b) and Bramson, Cox and Griffeath (1986). In Section 5 we prove Theorem 6 and then exploit the duality relationship between the voter model and coalescing random walk to prove Theorems 1–3 (for $d \geq 2$). Section 7 contains $d = 1$ proofs, using the work of D. Aldous (personal communication) and Arratia (1979). Section 8 concludes the paper by stating some extensions of our results to the multitype voter model.

A word about notation. We will use C to denote a finite positive constant whose value is unimportant; the value of C may change from line to line. We will also write $\varepsilon_N = \varepsilon_N(v_1, v_2, \dots)$ for quantities which depend on the variables $v_1, v_2, \dots$, but which tend to zero uniformly in these variables as $N \to \infty$; the value of ε_N may change from line to line.

2. Simple random walk on the torus. The goal of this section is to prove Theorem 4. As in the introduction, let $X_t^{(N)}$, $t \geq 0$, be rate 1 continuous time simple symmetric random walk on the torus $\Lambda(N)$, let $H^{(N)}$ be the hitting time of the origin and let $p_t^{(N)}(x, y) = P_x[X_t^{(N)} = y]$. Quantities without the superscript or subscript N will refer to random walk on $\mathbb{Z}^d$. Let F_N and G_N be the Laplace transforms

$$(2.1)\qquad \begin{aligned} F_N(x, \lambda) &= \int_0^\infty e^{-\lambda t} P_x[H^{(N)} \in dt],\\ G_N(x, \lambda) &= \int_0^\infty e^{-\lambda t} p_t^{(N)}(x, 0)\, dt, \end{aligned}$$

defined for $\lambda > 0$ and $x \in \Lambda(N)$. From the decomposition

$$G_N(x, \lambda) = \int_0^\infty e^{-\lambda t} \int_0^t P_x[H^{(N)} \in du]\, p_{t-u}^{(N)}(0,0)\, dt,$$

it is easy to derive the fundamental relation

$$(2.2)\qquad F_N(x, \lambda) = \frac{G_N(x, \lambda)}{G_N(0, \lambda)}, \qquad \lambda > 0,\ x \in \Lambda(N).$$

Now let ϕ be the characteristic function of discrete time random walk on $\mathbb{Z}^d$,

$$\phi(\theta) = \sum_x e^{ix\cdot\theta} p(0, x) = d^{-1} \sum_{j=1}^d \cos(\theta_j), \qquad \theta \in \mathbb{R}^d.$$

Then it is well known [see Montroll (1969) for instance] that

$$(2.3)\quad G_N(x, \lambda) = N^{-d} \sum_{y \in \Lambda(N)} \frac{e^{i2\pi x\cdot y/N}}{1 + \lambda - \phi(2\pi y/N)}, \qquad \lambda > 0,\ x \in \Lambda(N).$$

Actually, Montroll (1969) treats discrete time random walk, but only minor modifications are needed to handle the continuous time case. For a recent treatment of (2.3), with applications and extensions, see den Hollander and Kasteleyn (1982) [see also Flatto, Odlyzko and Wales (1985)].

As a warm up for what follows we present an unpublished result of F. Spitzer (personal communication) and give his proof. Let $e_1 = (1, 0, \dots, 0)$, let $\gamma_d = P_0[H < \infty]$, let δ_0 be the unit point mass at zero and let $\mathscr{E}(\alpha)$ denote the exponential distribution with parameter α.

THEOREM 7 (Spitzer). *Suppose $d \geq 3$ and $X_0^{(N)} = e_1$. Then as $N \to \infty$,*

$$\mathscr{L}\left(H^{(N)}/N^d\right) \Rightarrow \gamma_d \delta_0 + (1 - \gamma_d)\mathscr{E}(G^{-1}).$$

PROOF. A simple calculation shows that $G_N(e_1, \lambda) = (\lambda + 1)G_N(0, \lambda) - 1$, and so by (2.2),

$$E_{e_1}\left[e^{-\lambda H^{(N)}/N^d}\right] = \frac{(\lambda N^{-d} + 1)G_N(0, \lambda N^{-d}) - 1}{G_N(0, \lambda N^{-d})}.$$

By (2.3) we have

$$\begin{aligned} G_N(0, \lambda N^{-d}) &= \frac{1}{\lambda} + N^{-d} \sum_{\substack{y \in \Lambda(N) \\ y \neq 0}} \frac{1}{1 + \lambda N^{-d} - \phi(2\pi y/N)} \\ &\to \lambda^{-1} + \int_{[-1/2, 1/2]^d} \frac{d\theta}{1 - \phi(2\pi\theta)} \\ &= \lambda^{-1} + G, \end{aligned} \tag{2.4}$$

and the result follows, since $\gamma_d = (G - 1)/G$. □

We begin now our preparations for the proof of Theorem 4. It is natural to start with

$$E_x\left[e^{-\lambda H^{(N)}/s_N}\right] = \frac{G_N(x, \lambda/s_N)}{G_N(0, \lambda/s_N)} \tag{2.5}$$

and to make use of (2.3). For $d \geq 3$ the behavior of $G_N(0, \lambda/s_N)$ is given in (2.4). For $d = 2$, $s_N = N^2 \log N$ and

$$\begin{aligned} \frac{G_N(0, \lambda/N^2 \log N)}{\log N} &= \lambda^{-1} + \frac{1}{N^2 \log N} \sum_{\substack{y \in \Lambda(N) \\ y \neq 0}} \frac{1}{1 + \lambda/N^2 \log N - \phi(2\pi y/N)} \\ &\sim \lambda^{-1} + \frac{1}{N^2 \log N} \sum_{\substack{y \in \Lambda(N) \\ y \neq 0}} \frac{1}{1 - \phi(2\pi y/N)}, \end{aligned}$$

where we have used $f(N) \sim g(N)$ to mean $f(N)/g(N) \to 1$ as $N \to \infty$. The evaluation of the last sum is fairly standard [see Montroll (1969) or den Hollander and Kasteleyn (1982), for example]; it converges to $2/\pi$. Thus we have

$$\frac{G_N(0, \lambda/N^2 \log N)}{\log N} \to \lambda^{-1} + \frac{2}{\pi}, \qquad d = 2. \tag{2.6}$$

Unfortunately, it does not seem possible to deal with $G_N(x, \lambda/s_N)$ in a comparable fashion. So instead we write

$$G_N(x, \lambda/s_N) = \int_0^\infty \exp(-\lambda/s_N) p_t^{(N)}(x, 0)\, dt \tag{2.7}$$

and estimate this quantity by obtaining good estimates on $p_t^{(N)}(x, 0)$. The first step in doing so is the following:

PROPOSITION. *For $d \geq 2$, if $t_N \to \infty$, then*

$$\lim_{N \to \infty} \sup_{u \geq t_N N^2} \sup_{x \in \Lambda(N)} N^d \left| p_u^{(N)}(x, 0) - N^{-d} \right| = 0. \tag{2.8}$$

For $d = 2$, if $a_N \to \infty$ and $a_N = o(N)$ as $N \to \infty$, then there is a finite constant K such that

$$\limsup_{N \to \infty} \sup_{u \geq 1} \sup_{|x| \geq a_N,\, x \in \Lambda(N)} a_N^2 p_u^{(N)}(x, 0) \leq K. \tag{2.9}$$

Since $p_t^{(N)}(x, y)$ is doubly stochastic, $p_t^{(N)}(x, 0) \to N^{-d}$ as $t \to \infty$ for fixed N and x, but (2.8) provides some uniformity we will need later. The key to the proposition is a very precise expansion of $p_t(x, y)$. By Corollary 2.2.3 of Bhattacharya and Rao (1976), applied to $p_1(x, y)$, we have for $t = 1, 2, \dots,$

$$p_t(0, x) = \left(\frac{d}{2\pi t}\right)^{d/2} \exp\left(-\frac{d|x|^2}{2t}\right) \left\{1 + \sum_{r=1}^{d} t^{-r/2} B_r\left(\frac{x}{\sqrt{t}}\right)\right\} + e(x, t), \tag{2.10}$$

where each B_r is a polynomial (depending on d) of degree at most r and

$$t^{d/2} \sum_{x \in \mathbf{Z}^d} |e(x, t)| \to 0 \quad \text{as } t \to \infty. \tag{2.11}$$

PROOF OF (2.8). We first note that we may assume that t_N is a sequence of integers and that it suffices to prove that

$$\lim_{N \to \infty} \sup_{x \in \Lambda(N)} N^d \left| p_{t_N N^2}^{(N)}(x, 0) - N^{-d} \right| = 0. \tag{2.12}$$

For if $u \geq t_N N^2$ and $x \in \Lambda(N)$,

$$\begin{aligned} N^d \left| p_u^{(N)}(x, 0) - N^{-d} \right| &= N^d \left| \sum_y p_{u - t_N N^2}^{(N)}(x, y) \left[p_{t_N N^2}^{(N)}(y, 0) - N^{-d} \right] \right| \\ &\leq \sup_y N^d \left| p_{t_N N^2}^{(N)}(y, 0) - N^{-d} \right| \\ &\to 0 \end{aligned}$$

by (2.12).

Since $p_u^{(N)}(x,0) = \sum_{z\in\mathbf{Z}^d} p_u(x, Nz)$ it follows from (2.10) that

$$
\begin{aligned}
N^d p_u^{(N)}(x,0) &= N^d\left(\frac{d}{2\pi u}\right)^{d/2} \sum_z \exp\left(-\frac{d|x-Nz|^2}{2u}\right)\\
&\quad + N^d\left(\frac{d}{2\pi u}\right)^{d/2} \sum_z \exp\left(-\frac{d|x-Nz|^2}{2u}\right) \sum_{r=1}^{d} u^{-r/2} B_r\left(\frac{x-Nz}{\sqrt{u}}\right) \qquad (2.13)\\
&\quad + N^d \sum_z e(x-Nz, u).
\end{aligned}
$$

We let $u = t_N N^2$ and analyze this expansion in three parts.

(i) For the "main" term, fix $R > 0$, let $I = [-\frac{1}{2}, \frac{1}{2}]^d$ and write x' for x/N. Then

$$
\begin{aligned}
N^d&\left(\frac{d}{2\pi u}\right)^{d/2} \sum_z \exp\left(-\frac{d|x-Nz|^2}{2u}\right)\\
&= N^d\left(\frac{d}{2\pi u}\right)^{d/2} \sum_{|z|\le R} \exp\left(-\frac{N^2 d|x'-z|^2}{2u}\right)\\
&\quad + N^d\left(\frac{d}{2\pi u}\right)^{d/2} \sum_{|z|>R} \int_{I+z} \exp\left(-\frac{N^2 d|x'-z|^2}{2u}\right) dy.
\end{aligned}
$$

Then for some finite constant C, the first sum in the right-hand side above is majorized by

$$
\frac{CN^d R^d}{u^{d/2}} = \frac{CR^2}{t_N^{d/2}} \to 0
$$

as $N \to \infty$ for fixed R. For the second sum, observe that $x' \in I$ and we can choose κ_R such that $0 < \kappa_R < 1$, $\kappa_R \to 1$ as $R \to \infty$ and

$$
\kappa_R \le \frac{|y|}{|x'-z|} \le \frac{1}{\kappa_R}, \qquad |z| \ge R \text{ and } y \in I + z.
$$

Using this inequality we obtain

$$
\begin{aligned}
N^d&\left(\frac{d}{2\pi u}\right)^{d/2} \sum_{|z|>R} \int_{I+z} \exp\left(-\frac{dN^2|x'-z|^2}{2u}\right) dy\\
&\le N^d\left(\frac{d}{2\pi u}\right)^{d/2} \sum_{|z|\ge R-1} \int_{I+z} \exp\left(-\frac{dN^2\kappa_R^2|y|^2}{2u}\right) dy\\
&\le N^d\left(\frac{d}{2\pi u}\right)^{d/2} \left(\int_{-\infty}^{\infty} \exp\left(-\frac{dN^2\kappa_R^2 r^2}{2u}\right) dr\right)^d\\
&= \frac{1}{\kappa_R^d}\\
&\to 1
\end{aligned}
$$

as $R \to \infty$. On the other hand,

$$N^d\left(\frac{d}{2\pi u}\right)^{d/2} \sum_{|z|>R} \int_{I+z} \exp\left(-\frac{dN^2|x'-z|^2}{2u}\right) dy$$
$$\geq N^d\left(\frac{d}{2\pi u}\right)^{d/2} \int_{|y|>R+1} \exp\left(-\frac{dN^2|y|^2}{2\kappa_R^2 u}\right) dy$$
$$= \kappa_R^d (2\pi)^{-d/2} \int_{|y|>c_N} \exp(-|y|^2/2)\, dy,$$

where

$$c_N = (R+1)\sqrt{\frac{dN^2}{\kappa_R^2 u}} = (R+1)\sqrt{\frac{d}{\kappa_R^2 t_N}} \to 0$$

as $N \to \infty$ for fixed R. This proves that

$$\liminf_{N\to\infty} N^d\left(\frac{d}{2\pi u}\right)^{d/2} \sum_{|z|>R} \int_{I+z} \exp\left(-\frac{dN^2|x'-z|^2}{2u}\right) dy$$
$$\geq \kappa_R^d (2\pi)^{-d/2} \int_{\mathbb{R}^d} \exp\left(-\frac{|y|^2}{2}\right) dy$$
$$= \kappa_R^d.$$

Since $\kappa_R \to 1$ as $R \to \infty$, we have established

$$N^d\left(\frac{d}{2\pi t_N N^2}\right)^{d/2} \sum_z \exp\left(-\frac{d|x-Nz|^2}{2t_N N^2}\right) \to 1. \tag{2.14}$$

(ii) Since each B_r is a polynomial, the argument leading to (2.12) shows that

$$\limsup_{N\to\infty} \left\{ N^d\left(\frac{d}{2\pi t_N N^2}\right)^{d/2} \sum_z \exp\left(-\frac{d|x-Nz|^2}{2t_N N^2}\right) B_r\left(\frac{x-Nz}{\sqrt{t_N N^2}}\right)\right\} < \infty,$$

which implies

$$\lim_{N\to\infty} \left\{ N^d\left(\frac{d}{2\pi t_N N^2}\right)^{d/2} \sum_z \exp\left(-\frac{d|x-Nz|^2}{2t_N N^2}\right) \times\left[\left(t_N N^2\right)^{-r/2} B_r\left(\frac{x-Nz}{\sqrt{t_N N^2}}\right)\right]\right\} = 0. \tag{2.15}$$

(iii) It follows from (2.11) that

$$N^d \sum_z e\left(x-Nz, t_N N^2\right) \to 0. \tag{2.16}$$

By combining (2.13)–(2.16) we obtain (2.12). □

PROOF OF (2.9). It suffices to prove that

$$\limsup_{N\to\infty} \sup_{k\geq 1} \sup_{\substack{|x|\geq a_N \\ x\in\Lambda(N)}} a_N^2 p_k^{(N)}(x,0) < \infty, \tag{2.17}$$

where k denotes a positive integer. To see this, note that if $k \leq t \leq k+1$, then

$$\begin{aligned} p_{k+1}^{(N)}(x,0) &\geq p_t^{(N)}(x,0)p_{k+1-t}^{(N)}(0,0) \\ &\geq p_t^{(N)}(x,0)p_{k+1-t}(0,0) \\ &\geq c_0 p_t^{(N)}(x,0), \end{aligned}$$

where $c_0 = \inf\{p_s(0,0)\colon 0 \leq s \leq 1\} > 0$. The inequality $p_t^{(N)}(x,0) \leq c_0^{-1}p_{k+1}^{(N)}(x,0)$ and (2.17) imply (2.9).

Let $x' = x/N$ again. Then the main contribution to $a_N^2 p_k^{(N)}(x,0)$ [from an expansion like (2.13)] is

$$a_N^2 \sum_z \exp\left(-\frac{|x-Nz|^2/k}{\pi k}\right) = \frac{a_N^2}{\pi k}\left[\exp\left(-\frac{|x|^2}{k}\right) + \sum_{z\neq 0}\exp\left(-\frac{N^2|x'-z|^2}{k}\right)\right].$$

The first term above is bounded, since for some finite constant C,

$$\frac{a_N^2}{\pi k}\exp\left(-\frac{|x|^2}{k}\right) \leq \frac{a_N^2}{\pi k}\exp\left(-\frac{a_N^2}{k}\right) \leq C$$

for all k and N. For the second term, with $I = [-\frac{1}{2},\frac{1}{2}]^d$, we have

$$\begin{aligned} \frac{a_N^2}{\pi k}\sum_{z\neq 0}\exp\left(-\frac{N^2|x'-z|^2}{k}\right) &= \frac{a_N^2}{\pi k}\sum_{z\neq 0}\int_{z+I}\exp\left(-\frac{N^2|x'-z|^2}{k}\right)dy \\ &\leq \frac{a_N^2}{\pi k}\sum_{z\neq 0}\int_{z+I}\exp\left(-\frac{N^2\kappa_1^2|y|^2}{k}\right)dy \\ &\leq \frac{a_N^2}{\pi k}\int_{\mathbb{R}^2}\exp\left(-\frac{N^2\kappa_1^2|y|^2}{k}\right)dy \\ &= C\frac{a_N^2\kappa_1^2}{N^2} \\ &\to 0. \end{aligned}$$

This proves that

$$\limsup_{N\to\infty}\sup_{k\geq 1}\sup_{\substack{|x|\geq a_N \\ x\in\Lambda(N)}} a_N^2\sum_z \frac{e^{-|x-Nz|^2/k}}{\pi k} < \infty.$$

We omit the analysis of the error terms which completes the proof of (2.17). □

Proof of Theorem 4. It suffices to show that for any fixed $\lambda > 0$, uniformly in $|x| \geq a_N$,

$$E_x\left[\exp\left(-\frac{\lambda H^{(N)}}{s_N}\right)\right] \to \frac{1}{1+\lambda G} \quad \text{as } N \to \infty. \tag{2.18}$$

Since for each N the left side of (2.18) is a monotone function of λ and the right side is continuous in λ, it follows that (2.18) must hold uniformly in $\lambda \geq 0$. This implies that for any bounded continuous function f on $[0, \infty)$, uniformly in $|x| \geq a_N$,

$$E_x\left[f\left(H^{(N)}/s_N\right)\right] \to G^{-1}\int_0^\infty f(v)e^{-v/G}\,dv \quad \text{as } N \to \infty. \tag{2.19}$$

It now follows by approximation from (2.19) that for any fixed $t > 0$, uniformly in $|x| \geq a_N$,

$$P_x\left[H^{(N)}/s_N > t\right] \to e^{-t/G} \quad \text{as } N \to \infty. \tag{2.20}$$

Monotonicity implies (2.20) must hold uniformly in $t \geq 0$.

Turning to the proof of (2.18), recall that

$$E_x\left[e^{-\lambda H^{(N)}/s_N}\right] = \frac{G_N(x, \lambda/s_N)}{G_N(0, \lambda/s_N)}$$

and that by (2.4) and (2.6),

$$\frac{G_N(0, \lambda/N^2 \log N)}{\log N} \to \lambda^{-1} + G, \qquad d = 2,$$

$$G_N\left(0, \frac{\lambda}{N^d}\right) \to \lambda^{-1} + G, \qquad d \geq 3.$$

Thus it remains only to show [recall (2.7)] that uniformly in $|x| \geq a_N$,

$$\frac{1}{\log N}\int_0^\infty e^{-\lambda t/N^2 \log N} p_t^{(N)}(x, 0)\,dt \to \lambda^{-1}, \qquad d = 2 \tag{2.21}$$

and

$$\int_0^\infty e^{-\lambda t/N^d} p_t^{(N)}(x, 0)\,dt \to \lambda^{-1}, \qquad d \geq 3. \tag{2.22}$$

We begin with $d = 2$, assuming $a_N = o(N)$ and $a_N\sqrt{\log N}/N \to \infty$ as $N \to \infty$. Let $t_N \leq a_N\sqrt{\log N}/N$ such that $t_N \to \infty$ and $t_N = o(\log N)$, and break the integral in (2.21) into two parts. The first is

$$\begin{aligned}\frac{1}{\log N}\int_0^{t_N N^2} e^{-\lambda t/N^2 \log N} p_t^{(N)}(x, 0)\,dt &\leq \frac{1}{\log N}\left(1 + \int_1^{t_N N^2}\frac{K}{a_N^2}e^{-\lambda t/N^2 \log N}\,dt\right)\\ &\leq \frac{1}{\log N}\left[1 + \frac{KN^2 \log N}{\lambda a_N^2}\left(1 - e^{-\lambda t_N/\log N}\right)\right]\\ &\leq \frac{1}{\log N}\left(1 + \frac{KN^2 t_N}{a_N^2}\right)\\ &\to 0\end{aligned} \tag{2.23}$$

as $N \to \infty$, where we have used (2.9). The second part is

$$\begin{aligned}\frac{1}{\log N}\int_{t_N N^2}^{\infty} e^{-\lambda t/N^2 \log N} p_t^{(N)}(x,0)\,dt &\le \frac{1+o(1)}{N^2 \log N}\int_{t_N N^2}^{\infty} e^{-\lambda t/N^2 \log N}\,dt\\ &= \frac{1+o(1)}{\lambda}\exp\left(-\frac{\lambda t_N}{\log N}\right)\\ &\to \lambda^{-1}\end{aligned}$$

as $N \to \infty$, where we have used (2.8). This finishes the proof of (2.21).

An additional estimate we need for proving (2.22) is

$$P\left[|X_t| \ge t^{1/2}\log t\right] \le C/t^2, \qquad t \ge 0.$$

This is easily proved using exponential type estimates. Now for any finite set $\Gamma \subset \mathbb{Z}^d$ we have

$$\begin{aligned}p_t^{(N)}(x,0) &= \sum_{z\in\mathbb{Z}^d} p_t(x, Nz)\\ &\le \left|\Gamma \cap (x + N\mathbb{Z}^d)\right| p_t(0,0) + \sum_{z\notin\Gamma} p_t(0, x + Nz).\end{aligned}$$

Since $p_t(0,0) \le C/t^{d/2}$, the choice $\Gamma = [-t^{1/2}\log t, t^{1/2}\log t]^d$ above yields

$$\begin{aligned}p_t^{(N)}(0,x) &\le \frac{C}{t^{d/2}}\left(\frac{t^{1/2}\log t}{N}\right)^d + P\left[|X_t| \ge t^{1/2}\right]\\ &\le C\left\{\left(\frac{\log t}{N}\right)^d + \frac{1}{t^2}\right\}.\end{aligned}$$

Now break the integral in (2.22) into two parts. The first is

$$\begin{aligned}&\int_0^{N^2\log N} e^{-\lambda t/N^d} p_t^{(N)}(x,0)\,dt\\ &\quad\le \int_0^{N^2\log N} p_t^{(N)}(x,0)\,dt\\ &\quad\le T \sup_{|x|\ge a_N}\sup_{0\le u\le T} p_t^{(N)}(x,0) + C\int_T^{N^2\log N}\left(\frac{(\log N)^d}{N^d} + \frac{1}{t^2}\right)dt\\ &\quad\to \frac{C}{T}\end{aligned}$$

as $N \to \infty$ for fixed T. Now let $T \to \infty$ to obtain

$$\lim_{N\to\infty}\int_0^{N^2\log N} e^{-\lambda t/N^d} p_t^{(N)}(x,0)\,dt = 0.$$

Using (2.8) the second integral is

$$\int_{N^2 \log N}^{\infty} e^{-\lambda t/N^d} p_t^{(N)}(0, x)\, dt = (1 + o(1)) N^{-d} \int_{N^2 \log N}^{\infty} e^{-\lambda t/N^d}\, dt$$
$$\to \lambda^{-1}$$

as $N \to \infty$. This completes the proof of (2.22). □

3. Coalescing random walk on the torus—probability estimates. We assume throughout this section that $d \ge 2$. The $n = 2$ case of Theorem 5 is covered by (1.9), so we turn to the analysis for $n \ge 3$. The major step in the argument is establishing that

$$P\Big[\big|\xi_{ts_N}^{(N)}(A^{(N)})\big| = n\Big] \to \exp\Big[-2t\binom{n}{2}\Big/G\Big] \tag{3.1}$$

uniformly in $A^{(N)}$ and t as $N \to \infty$. To obtain (3.1) we will follow the approach of Cox and Griffeath (1986), keeping as close as possible to the notation used there. Define

$$\begin{aligned}
\tau^{(N)}(i, j) &= \inf\Big\{t \ge 0\colon \big|\xi_t^{(N)}(\{x_i, x_j\})\big| = 1\Big\},\\
\bar{\tau}^{(N)} &= \min_{i \ne j} \tau^{(N)}(i, j),\\
H_t^{(N)}(i, j) &= \big\{\tau^{(N)}(i, j) \le ts_N\big\},\\
F_t^{(N)}(i, j) &= \big\{\bar{\tau}^{(N)} = \tau^{(N)}(i, j) \le ts_N\big\},\\
q^{(N)}(t) &= P\big[\bar{\tau}^{(N)} \le ts_N\big].
\end{aligned}$$

Using this notation (3.1) is equivalent to

$$q^{(N)}(t) \to 1 - \exp\Big[-2t\binom{n}{2}\Big/G\Big] \tag{3.2}$$

and (1.9) is equivalent to

$$P\big(H_t^{(N)}(i, j)\big) \to 1 - e^{-2t/G}. \tag{3.3}$$

By examining which pair of particles hits first, we have, for $i \ne j$,

$$\begin{aligned}
P\big(H_t^{(N)}(i, j)\big) &= P\big(F_t^{(N)}(i, j)\big)\\
&\quad + \sum_{\{k, l\} \ne \{i, j\}} \sum_{y_\alpha, y_\beta} \int_0^{ts_N} P\big[\bar{\tau}^{(N)} = \tau^{(N)}(k, l) \in du, \xi_u^{(N)}(x_i) = y_\alpha,\\
&\qquad \xi_u^{(N)}(x_j) = y_\beta\big] P\Big[\big|\xi_{ts_N - u}^{(N)}(\{y_\alpha, y_\beta\})\big| = 1\Big].
\end{aligned} \tag{3.4}$$

It will follow from (3.7) and (3.8) that

$$\begin{aligned}
&\int_0^{Ts_N} P\Big[\bar{\tau}^{(N)} = \tau^{(N)}(k, l) \in du, \big|\xi_u^{(N)}(x_i) - \xi_u^{(N)}(x_j)\big| \le a_N\Big]\\
&\qquad = \varepsilon_N \to 0,
\end{aligned} \tag{3.5}$$

so we may assume in (3.4) that $|y_\alpha - y_\beta| \geq a_N$. In this case

$$P\Big[\big|\xi^{(N)}_{ts_N-u}(\{y_\alpha, y_\beta\})\big| = 1\Big] = P_y\big[H^{(N)} \leq 2(ts_N - u)\big]$$
$$= 1 - \exp\left[-2\left(t - \frac{u}{s_N}\right)\Big/G\right] + \varepsilon_N$$

by Theorem 4, where $y = y_\alpha - y_\beta$. Consequently

$$\sum_{y_\alpha, y_\beta} \int_0^{ts_N} P\big[\bar{\tau}^{(N)} = \tau^{(N)}(k,l) \in du,\ \xi_u^{(N)}(x_i) = y_\alpha,\ \xi_u^{(N)}(x_j) = y_\beta\big]$$
$$\times P\Big[\big|\xi^{(N)}_{ts_N-u}(\{y_\alpha, y_\beta\})\big| = 1\Big]$$
$$= \int_0^{ts_N} P\big[\bar{\tau}^{(N)} = \tau^{(N)}(k,l) \in du\big]\left(1 - \exp\left[-2\left(t - \frac{u}{s_N}\right)\Big/G\right]\right) + \varepsilon_N$$
$$= \int_0^t P\big(F_u^{(N)}(k,l)\big)e^{-2(t-u)/G}\,du + \varepsilon_N,$$

the last equality from integration by parts and a change of variables.

Combining this last result with (3.3) and (3.4) we see that

$$1 - e^{-2t/G} = P\big(F_t^{(N)}(i,j)\big) + \frac{2}{G}\sum_{\{k,l\}\neq\{i,j\}} \int_0^t P\big(F_u^{(N)}(k,l)\big)e^{-2(t-u)/G}\,du + \varepsilon_N.$$

Summation over i, j leads to

$$\binom{n}{2}\left(1 - e^{-2t/G}\right)$$
$$= q^{(N)}(t) + \frac{2}{G}e^{-2t/G}\left[\binom{n}{2} - 1\right]\int_0^t q^{(N)}(u)e^{2u/G}\,du + \varepsilon_N. \tag{3.6}$$

It now follows [see the proof of Lemma 2 of Cox and Griffeath (1986)] that as $N \to \infty$, $q^{(N)}(t) \to$ the solution of

$$\binom{n}{2}\left(1 - e^{-2t/G}\right) = q(t) + \frac{2}{G}e^{-2t/G}\left[\binom{n}{2} - 1\right]\int_0^t q(u)e^{2u/G}\,du,$$

that is,

$$q^{(N)}(t) \to q(t) \equiv 1 - \exp\left(-2t\binom{n}{2}\Big/G\right)$$

uniformly in $A^{(N)}$ and t. This completes the proof of (3.2) except for the justification of (3.5), which we will now carry out.

Proof of (3.5). There are two cases to consider. Let $X_u^{(N)}(x)$, $x \in \Lambda(N)$, be independent random walks, $X_0^{(N)}(x) = x$. We must show

$$\int_0^{Ts_N} P\left[\bar{\tau}^{(N)} = \tau^{(N)}(1,2) \in du, \left|\xi_u^{(N)}(x_3) - \xi_u^{(N)}(x_4)\right| \le a_N\right] \to 0, \tag{3.7}$$

$$\int_0^{Ts_N} P\left[\bar{\tau}^{(N)} = \tau^{(N)}(1,2) \in du, \left|\xi_u^{(N)}(x_3) - \xi_u^{(N)}(x_1)\right| \le a_N\right] \to 0. \tag{3.8}$$

We will prove only (3.8) as the proof of (3.7) is similar (and slightly easier). If $t_N N^2 \le Ts_N$, then by using the independent random walks $X_u^{(N)}(x)$ we can write

$$\begin{aligned}
&\int_0^{Ts_N} P\left[\bar{\tau}^{(N)} = \tau^{(N)}(1,2) \in du, \left|\xi_u^{(N)}(x_3) - \xi_u^{(N)}(x_1)\right| \le a_N\right] \\
&\quad \le P\left[\tau^{(N)}(1,2) \le t_N N^2\right] \\
&\qquad + \int_{t_N N^2}^{Ts_N} P\left[\tau^{(N)}(1,2) \in du, \left|X_u^{(N)}(x_3) - X_u^{(N)}(x_1)\right| \le a_N\right] \qquad (3.9) \\
&\quad = P\left[\tau^{(N)}(1,2) \le t_N N^2\right] + \sum_y \int_{t_N N^2}^{Ts_N} P\left[\tau^{(N)}(1,2) \in du,\ X_u^{(N)}(x_1) = y\right] \\
&\qquad\qquad \times P\left[\left|X_u^{(N)}(x_3) - y\right| \le a_N\right],
\end{aligned}$$

where we have also used $\tau^{(N)}(1,2)$ to denote the first time $X_u^{(N)}(x_1) = X_u^{(N)}(x_2)$.

Choose t_N such that $t_N \to \infty$, $t_N/\log N \to 0$. Then the right-hand side of (3.9) is majorized by

$$\begin{aligned}
&P_{x_1 - x_2}\left[H^{(N)} \le 2t_N N^2\right] + C\frac{a_N^d}{N^d} P\left[\tau^{(N)}(1,2) \in \left[t_N N^2, Ts_N\right]\right] \\
&\quad \le P_{x_1 - x_2}\left[\frac{H^{(N)}}{s_N} \le \frac{2t_N N^2}{s_N}\right] + C\frac{a_N^d}{N^d} \\
&\quad \to 0
\end{aligned}$$

using Theorem 4 and (2.8), since $t_N N^2/s_N \to 0$. □

With (3.2) established we can finish the proof of Theorem 5 by induction. The induction hypothesis is that for a_N satisfying the assumptions of Theorem 4, uniformly for $t \in [0, T]$ and $A^{(N)} = \{x_1, \dots, x_n\} \subset \Lambda(N)$ such that $|x_i - x_j| \ge a_N$ for $i \ne j$,

$$P\left[\left|\xi_{ts_N}^{(N)}(A^{(N)})\right| = k\right] \to q_{n,k}(t), \qquad 1 \le k \le n.$$

The case $n = 2$ is covered by (1.9) and the $n = k$ case is covered by (3.2). The induction step is to prove that uniformly for $t \in [0, T]$ and $B^{(N)} = \{y_1, \dots, y_{n+1}\} \subset \Lambda(N)$ such that $|y_i - y_j| \ge a_N$ for $i \ne j$,

$$\begin{aligned}
P\left[\left|\xi_{ts_N}^{(N)}(B^{(N)})\right| = k\right] &\to \binom{n+1}{2}\frac{2}{G}\int_0^t \exp\left(-2u\binom{n+1}{2}\Big/G\right) q_{n,k}(t-u)\,du \\
&= q_{n+1,k}(t), \qquad 1 \le k \le n.
\end{aligned}$$

To prove this let $\bar{\sigma}^{(N)} = \inf\{t \geq 0: |\xi_t^{(N)}(B^{(N)})| = n\}$ and fix $k \leq n$. Then

$$P\left[\left|\xi_{ts_N}^{(N)}(B^{(N)})\right| = k\right]$$
$$= \sum_{C^{(N)}} \int_0^{ts_N} P\left[\bar{\sigma}^{(N)} \in du,\ \xi_u^{(N)}(B^{(N)}) = C^{(N)}\right] P\left[\left|\xi_{ts_N - u}^{(N)}(C^{(N)})\right| = k\right],$$

where $C^{(N)} = \{z_1, \ldots, z_n\} \subset \Lambda(N)$. It is a consequence of (3.5) that

$$\int_0^{Ts_N} P\left[\bar{\sigma}^{(N)} \in du,\ \xi_u^{(N)}(B^{(N)}) = \{z_1, \ldots, z_n\} \text{ and } |z_i - z_j| \leq a_N \text{ for some } i \neq j\right] \to 0,$$

and so by using the induction hypothesis we obtain

$$P\left[\left|\xi_{ts_N}^{(N)}(B^{(N)})\right| = k\right]$$
$$= \sum_{C^{(N)}} \int_0^{ts_N} P\left[\bar{\sigma}^{(N)} \in du,\ \xi_u^{(N)}(B^{(N)}) = C^{(N)}\right] \left(q_{n,k}\left(t - \frac{u}{s_N}\right) + \varepsilon_N\right) + \varepsilon_N$$
$$= \int_0^{ts_N} P[\bar{\sigma}^{(N)} \in du]\, q_{n,k}\left(t - \frac{u}{s_N}\right) + \varepsilon_N$$
$$= \binom{n+1}{2} \frac{2}{G} \int_0^t \exp\left(-2u\binom{n+1}{2}\Big/G\right) q_{n,k}(t-u)\, du + \varepsilon_N$$
$$\to \binom{n+1}{2} \frac{2}{G} \int_0^t \exp\left(-2u\binom{n+1}{2}\Big/G\right) q_{n,k}(t-u)\, du$$

as required. This completes the proof of Theorem 5.

In preparation for the proof of Theorem 2 we will establish another coalescing random walk result, this time for random walks that start moving at different times. For $t_1 < \cdots < t_k$ and $A_i \subset \Lambda(N)$ let $\xi_t^{(N)}(A_1, t_1; \ldots; A_k, t_k)$ denote a coalescing random walk system in which random walks start from each point of A_i at time t_i (they are *frozen* until this time) and then execute coalescing random walk motion. These systems were used in Cox and Griffeath (1983). For $t > t_k$ and positive integers m, n_i define

$$q_{n_1; m}(t_1; t) = q_{n_1, m}(t - t_1),$$
$$q_{n_1, n_2, \ldots, n_k; m}(t_1, t_2, \ldots, t_k; t) = \sum_{l_1} \sum_{l_2} \cdots \sum_{l_{k-1}} q_{n_1, l_1}(t_2 - t_1) q_{n_2 + l_1, l_2}(t_3 - t_2)$$
$$\times q_{n_{k-1} + l_{k-2}, l_{k-1}}(t_k - t_{k-1}) q_{n_k + l_{k-1}, m}(t - t_k).$$

It is straightforward to check that

$$\begin{aligned} & q_{n_1, n_2, \ldots, n_k; m}(t_1, t_2, \ldots, t_k; t) \\ & = \sum_l q_{n_1, n_2, \ldots, n_{k-1}; l}(t_1, t_2, \ldots, t_{k-1}; t_k) q_{n_k + l, m}(t - t_k) \\ & = \sum_l q_{n_1, l}(t_2 - t_1) q_{l + n_2, n_3, \ldots, n_k; m}(t_2, \ldots, t_k; t). \end{aligned} \tag{3.10}$$

THEOREM 8. *Assume* $d \geq 2$, *fix* $T > 0$, $k > 0$, $n_1, \dots, n_k$, *let* $a_N \to \infty$, $a_N = o(N)$ *as* $N \to \infty$. *Then for fixed* $0 \leq t_1 < \cdots < t_k < t$, *uniformly for* $A_i = \{x_i^j,\ j = 1, \dots, n_i\} \subset \Lambda(N)$ *such that* $|x_i^\alpha - x_i^\beta| \geq a_N$ *for all* i *and all* $\alpha \neq \beta$,

$$P\Big[\big|\xi_{ts_N}^{(N)}(A_1, t_1 s_N; \cdots; A_k, t_k s_N)\big| = m\Big] \to q_{n_1, n_2, \dots, n_k; m}(2t_1/G, \dots, 2t_k/G; 2t/G). \tag{3.11}$$

PROOF. The $k = 1$ result follows from Theorem 5, so we assume now that $k \geq 2$ and proceed by induction. The idea is to run the system until time $t_2 s_N$ and look at $\xi_{ts_N}^{(N)}(A_1, t_1 s_N) = \{y_1, \dots, y_l\}$. By constructing independent random walks and applying the proposition of Section 2 it is easy to see that

$$P\Big[\exists\ y_\alpha, y_\beta \in \xi_{ts_N}^{(N)}(A_1, t_1 s_N),\ y_\alpha \neq y_\beta \text{ and } |y_\alpha - y_\beta| \leq a_N\Big] \leq C a_N^d / N^d,$$

$$P\Big[\exists\ y_\alpha \in \xi_{ts_N}^{(N)}(A_1, t_1 s_N) \text{ and } x_\beta \in A_2 \text{ with } |y_\alpha - x_\beta| \leq a_N\Big] \leq C a_N^d / N^d.$$

Thus

$$\begin{aligned} P\Big[|\xi_{ts_N}^{(N)}| = m\Big] = \varepsilon_N + \sum_{\ell} \sum_{y_1, \dots, y_\ell} & P\big[\xi_{ts_N}^{(N)}(A_1, ts_N) = \{y_1, \dots, y_\ell\}\big] \\ & \times P\Big[\big|\xi_{ts_N - t_2 s_N}^{(N)}(A_2 \cup \{y_1, \dots, y_\ell\}, 0; A_3, t_3 s_N - t_2 s_N; \\ & \qquad \cdots; A_k, t_k s_N - t_2 s_N)\big| = m\Big] \end{aligned}$$

(where the sum on the y_i is over $|y_\alpha - y_\beta| \geq a_N$ and $|y_\alpha - x_\beta| \geq a_N$, all $x_\beta \in A_2$)

$$\begin{aligned} &\to \sum q_{n_1, \ell}(2(t_2 - t_1)/G) \\ &\quad \times q_{n_2 + \ell, n_3, \dots, n_k; m}(0, 2(t_3 - t_2)/G, \dots, 2(t_k - t_2)/G; 2(t - t_2)/G) \\ &= q_{n_1, n_2, \dots, n_k; m}(2t_1/G, \dots, 2t_k/G; 2t/G) \end{aligned}$$

by using Theorem 5, the induction hypothesis and (3.10). □

4. Coalescing random walk on the torus—expectation estimates. In order to prove Theorem 6 we need some control over the number of particles left in the coalescing random walk system $\xi_t^{(N)} = \xi_t^{(N)}(\Lambda(N))$ (we will use this abbreviation throughout this section). The following result gives us this control. It is similar to Theorem 1 of Bramson and Griffeath (1980b) and the proposition in the introduction of Bramson, Cox and Griffeath (1986), although the proof here is easier. First define $g_N(t)$ by

$$g_N(t) = \begin{cases} N/\sqrt{t}, & d = 1, \\ N^2 \log(1 + t)/t, & d = 2, \\ N^d/t, & d \geq 3. \end{cases}$$

PROPOSITION. *There are finite constants c_d such that*

$$E\left[|\xi_t^{(N)}|\right] \le c_d \max(1, g_N(t)), \qquad t > 0,\ N = 2, 4, \dots. \tag{4.1}$$

We will prove this proposition via a series of lemmas. The first is:

LEMMA. *If $B \subset A \subset \Lambda(N)$ and $h_s(A) = \min_{x,\, y \in \Lambda(N)} P_{x-y}[H^{(N)} \le s]$, then*

$$E\left[\left|\xi_t^{(N)}(B)\right|\right] \le |B| - (|B| - 1)h_s(A). \tag{4.2}$$

PROOF. We may assume that $|B| \ge 1$, so fix $x_0 \in B$ and define

$$Z_s = \sum_{x \in B \setminus \{x_0\}} 1\left(\xi_t^{(N)}(x) = \xi_t^{(N)}(x_0)\right),$$

the number of walks which coalesce with the walk starting at x_0. Observe that

$$\left|\xi_s^{(N)}(B)\right| \le |B| - Z_s.$$

To get (4.2) take expectations and use $E[Z_s] \ge (|B| - 1)h_s(A)$. □

For the next step, let $\lfloor t \rfloor$ denote the greatest integer less than or equal to t.

LEMMA. *If $t \le r \le r + s \le 2t$, $\frac{1}{2}E[|\xi_{2t}^{(N)}|] \ge 4^{-d}E[|\xi_t^{(N)}|] \ge 2^d$ and A_t is a cube in $\Lambda(N)$ of side $\lfloor 8N/E[|\xi_t^{(N)}|]^{1/d} \rfloor$, then*

$$E\left[|\xi_{r+s}^{(N)}|\right] \le E\left[|\xi_r^{(N)}|\right]\left(1 - \tfrac{1}{2}h_s(A_t)\right) \le E\left[|\xi_r^{(N)}|\right]\exp\left(-\tfrac{1}{2}h_s(A_t)\right). \tag{4.3}$$

PROOF. Let B_i, $1 \le i \le n(t)$, be disjoint cubes covering $\Lambda(N)$, each B_i no larger than A_t, with

$$\begin{aligned} n(t) &\le \left[\frac{NE\left[|\xi_t^{(N)}|\right]^{1/d}}{8N} + 1\right]^d \\ &\le \left[\frac{E\left[|\xi_t^{(N)}|\right]^{1/d}}{4}\right]^d \\ &= \frac{E\left[|\xi_t^{(N)}|\right]}{4^d} \\ &\le \frac{1}{2}E\left[|\xi_{2t}^{(N)}|\right] \\ &\le \frac{1}{2}E\left[|\xi_r^{(N)}|\right]. \end{aligned}$$

If we ignore the coalescence of particles starting in different B_i's, then the Markov property implies

$$E\left[|\xi_{r+s}^{(N)}|\right] \le \sum_{C_i \subset B_i,\, 1 \le i \le n(t)} P\left[\xi_r^{(N)} \cap B_i = C_i\right] \sum_i E\left[\left|\xi_s^{(N)}(C_i)\right|\right]. \tag{4.4}$$

Using the inequalities (4.2) and (4.3) and writing h_s for $h_s(A_t)$ we have

$$\begin{aligned}\sum_i E\left[\left|\xi_s^{(N)}(C_i)\right|\right] &\le \sum_i \left[|C_i| - (|C_i| - 1)h_s\right]\\ &= (1 - h_s)\sum_i |C_i| + h_s n(t)\\ &\le (1 - h_s)\sum_i |C_i| + \tfrac{1}{2}h_s E\left[|\xi_r^{(N)}|\right].\end{aligned}$$

Using this estimate in (4.4) we obtain

$$E\left[|\xi_{r+s}^{(N)}|\right] \le \left(1 - \tfrac{1}{2}h_s\right)E\left[|\xi_r^{(N)}|\right],$$

as required. □

LEMMA. *If* $f_N(t) = E[|\xi_t^{(N)}|]/g_N(t)$, *then there exist finite constants* M_d *such that*

(4.5) $$f_N(t) \le M_d, \qquad 0 \le t \le 4,\ N = 2,4,\dots,$$

(4.6) $$f_N(2t) \le \max\{M_d, f_N(t)\}, \qquad t \ge 0,\ N = 2,3,\dots.$$

PROOF. If $t \le 4$, then since $|\xi_t^{(N)}| \le N^d$, (4.5) holds with $M_d = 4$. Now if we cannot apply (4.3), then either

$$E\left[|\xi_t^{(N)}|\right] \le 8^d$$

or

$$E\left[|\xi_{2t}^{(N)}|\right] \le 2 \cdot 4^{-d} E\left[|\xi_t^{(N)}|\right],$$

and in either case, for all d it is easy to see that (4.6) holds with $M_d = 8^d$. So we may assume that (4.3) can be applied, in which case iteration of (4.3) gives

(4.7) $$E\left[|\xi_{2t}^{(N)}|\right] \le E\left[|\xi_t^{(N)}|\right]\exp\left(-\frac{1}{2}\left\lfloor\frac{t}{s}\right\rfloor h_s(A_t)\right).$$

To employ (4.7) effectively we recall from Lemma 5 of Bramson and Griffeath (1980b) that if B is a square of side $b \ge 8$, then there are positive constants α_d such that

$$h_{b^2}(B) \ge \begin{cases}\alpha_1, & d = 1,\\ \alpha_2/\log b, & d = 2,\\ \alpha_3/b^{d-2}, & d \ge 3.\end{cases}$$

We now let s depend on t by setting s to be the square of the side of A_t, i.e., $s = s_t = (8N/E[|\xi_t^{(N)}|]^{1/d})^2$. We may assume that $s_t \le t/2$, else it is easy to check that $f_N(t) \le 128^{d/2}$, $t \ge 2$. With this choice we have

(4.8) $$\left\lfloor\frac{t}{s_t}\right\rfloor \ge \frac{t}{2s_t} \ge \frac{tE\left[|\xi_t^{(N)}|\right]^{2/d}}{128N^2}$$

and

$$(4.9)\qquad h_{s_t}(A_t) \geq \begin{cases} \alpha_1, & d = 1, \\ \alpha_2/\log\left[\dfrac{8N}{E\left[|\xi_t^{(N)}|\right]^{1/2}}\right], & d = 2, \\ \alpha_d\Big/\left[\dfrac{8N}{E\left[|\xi_t^{(N)}|\right]^{1/2d}}\right]^{d-2}, & d \geq 3. \end{cases}$$

We now consider the cases $d = 1$, $d = 2$ and $d \geq 3$ separately.

$d = 1$: Utilizing (4.8), (4.9) and (4.10) we have

$$\begin{aligned} f_N(2t) &= \frac{\sqrt{2t}\,E\left[|\xi_{2t}^{(N)}|\right]}{N} \\ &\leq \frac{\sqrt{2t}\,E\left[|\xi_t^{(N)}|\right]}{N}\exp\left[-\frac{1}{2}\frac{tE\left[|\xi_t^{(N)}|\right]^2}{128N^2}\alpha_1\right] \\ &= f_N(t)\exp\left[\log\sqrt{2} - \frac{\alpha_1}{256}f_N^2(t)\right] \end{aligned}$$

and so (4.6) holds with $M_1 = \max\{\sqrt{128}, \sqrt{128\log 2/\alpha_1}\}$.

$d = 2$: As with $d = 1$ we have

$$\begin{aligned} f_N(2t) &= \frac{2tE\left[|\xi_{2t}^{(N)}|\right]}{N^2\log 2t} \\ &\leq \frac{2tE\left[|\xi_t^{(N)}|\right]}{N^2\log 2t}\exp\left[-\frac{1}{2}\frac{tE\left[|\xi_t^{(N)}|\right]}{128N^2}\frac{\alpha_2}{\log\left(8N/\left(E\left[|\xi_t^{(N)}|\right]^{1/2}\right)\right)}\right] \\ &= f_N(t)\left(\frac{\log(t)}{\log(2t)}\right)^2\exp\left[\log 2 - \frac{\alpha_2}{256}f_N(t)\frac{\log t}{\log\left(8\sqrt{t}/\sqrt{\log t}\right) - \frac{1}{2}\log f_N(t)}\right] \\ &\leq f_N(t)\exp\left[\log 2 - \frac{\alpha_2}{256}f_N(t)\right], \end{aligned}$$

unless

$$\frac{\log t}{\log\left(8\sqrt{t}/\sqrt{\log t}\right) - \frac{1}{2}\log f_N(t)} \leq 1.$$

Since the denominator is positive ($f_N(t) \leq t/\log t$) this can happen only if

$$\begin{aligned} \log f_N(t) &\leq 4\left(\log\frac{8\sqrt{t}}{\sqrt{\log t}} - \log t\right) \\ &\leq 4\log 8, \end{aligned}$$

i.e., $f_N(t) \le 8^2$. Putting all of the pieces together we have (4.7) with $M_2 = \max\{128, 256\log 2/\alpha_2\}$.

$d \ge 3$: As above,

$$\begin{aligned} f_N(2t) &= \frac{2tE\left[|\xi_{2t}^{(N)}|\right]}{N^d} \\ &\le \frac{2tE\left[|\xi_t^{(N)}|\right]}{N^d}\exp\left[-\frac{1}{2}\frac{tE\left[|\xi_t^{(N)}|\right]^{2/d}}{128N^2}\alpha_d\left(\frac{E\left[|\xi_t^{(N)}|\right]^{1/d}}{8N}\right)^{d-2}\right] \\ &= f_N(t)\exp\left[\log 2 - \frac{\alpha_d}{4\cdot 8^d}f_N(t)\right] \end{aligned}$$

and so (4.7) holds with $M_1 = \max\{128^{d/2}, 4\cdot 8^d\log 2/\alpha_d\}$. □

The proof of the Proposition is now almost immediate, and is exactly the same as the last paragraph in the proof of Lemma 4 in Bramson, Cox and Griffeath (1986).

5. Proof of Theorem 6, $d \ge 2$. We start with the proof of (1.11). Fix $t > 0$, $j \ge 1$ and a_N as in Theorem 4. Now fix $n \ge 2$ and select $A^{(N)} = \{x_1, \dots, x_n\} \subset \Lambda(N)$, $|x_\alpha - x_\beta| \ge a_N$ for $\alpha \ne \beta$. Then since $\xi_t^{(N)}(A^{(N)}) \subset \xi_t^{(N)}(\Lambda(N))$,

$$\begin{aligned} P\left[\left|\xi_{ts_N}^{(N)}(\Lambda(N))\right| \le j\right] &\le P\left[\left|\xi_{ts_N}^{(N)}(A^{(N)})\right| \le j\right] \\ &\to \sum_{k=1}^{j} q_{n,k}(e^{-2t/G}) \end{aligned} \tag{5.1}$$

as $N \to \infty$ by Theorem 5. Letting $n \to \infty$ we obtain

$$\limsup_{N\to\infty} P\left[\left|\xi_{ts_N}^{(N)}(\Lambda(N))\right| \le j\right] \le \sum_{k=1}^{j} q_{\infty,k}(e^{-2t/G}).$$

For the reverse inequality fix M (large) and $\delta_1, \delta_2 > 0$ (small). By Chebyshev and the proposition of Section 4,

$$\begin{aligned} P\left[\left|\xi_{\delta_1 s_N}^{(N)}(\Lambda(N))\right| \ge M\right] &\le E\left[\left|\xi_{\delta_1 s_N}^{(N)}(\Lambda(N))\right|\right]/M \\ &\le c_d/\delta_1 M. \end{aligned}$$

It follows from (2.8) and the usual construction with independent random walks that, uniformly in $k \le M$ and $\{y_1, \dots, y_k\} \subset \Lambda(N)$,

$$\begin{aligned} &P\left[\xi_{\delta_1 s_N}^{(N)}(\Lambda(N)) = \{y_1, \dots, y_k\}, \exists\, z_1 \ne z_2 \in \xi_{\delta_2 s_N}^{(N)}(\Lambda(N)), |z_1 - z_2| \le a_N\right] \\ &\quad = \varepsilon_N \to 0. \end{aligned}$$

Combining these remarks with Theorem 5 applied to $\xi_{(t-\delta_2)s_N}^{(N)}(\xi_{\delta_2 s_N}^{(N)}(\Lambda(N)))$ and letting

$$A_\ell(z_1, \dots, z_\ell) = \left\{\xi_{\delta_2 s_N}^{(N)}(\Lambda(N)) = \{z_1, \dots, z_\ell\}, |z_\alpha - z_\beta| \ge a_N \text{ for } \alpha \ne \beta\right\},$$

we have

$$\begin{aligned}
P\Big[\big|\xi_{ts_N}^{(N)}(\Lambda(N))\big| \le j\Big] &\ge \sum_{\ell=1}^{M} P\Big[\big|\xi_{(t-\delta_2)s_N}^{(N)}(\{z_1,\dots,z_\ell\})\big| \le j \big| A_\ell(z_1,\dots,z_\ell)\Big] \\
&\quad \times P\Big[A_\ell(z_1,\dots,z_\ell) \big| \, \big|\xi_{\delta_1 s_N}^{(N)}(\Lambda(N))\big| \le M\Big] \\
&\quad \times P\Big[\big|\xi_{\delta_1 s_N}^{(N)}(\Lambda(N))\big| \le M\Big] \\
&\ge \left(1 - \frac{c_d}{\delta_1 M}\right) \sum_{\ell=1}^{M} \left(\sum_{k=1}^{j} q_{\ell,k}\big(e^{-2(t-\delta_2)/G}\big) + \varepsilon_N \right) \\
&\quad \times P\Big[A_\ell(z_1,\dots,z_\ell) \big| \, \big|\xi_{\delta_1 s_N}^{(N)}(\Lambda(N))\big| \le M\Big] \\
&\ge \left(1 - \frac{c_d}{\delta_1 M}\right) \left(\sum_{k=1}^{j} q_{\infty,k}\big(e^{-2(t-\delta_2)/G}\big) \right) + \varepsilon_N .
\end{aligned}$$

If we first let $N \to \infty$, then $M \to \infty$ and then δ_1 and $\delta_2 \to 0$, we obtain

$$\liminf_{N\to\infty} P\Big[\big|\xi_{ts_N}^{(N)}(\Lambda(N))\big| \le j\Big] \ge \sum_{k=1}^{j} q_{\infty,k}\big(e^{-2t/G}\big).$$

This inequality and (5.1) prove

$$\lim_{N\to\infty} P\Big[\big|\xi_{ts_N}^{(N)}(\Lambda(N))\big| = j\Big] = q_{\infty,j}\big(e^{-2t/G}\big),$$

which is enough to prove the weak convergence in (1.11).

The moment convergence in (1.11) follows from weak convergence provided the sequence $\sigma_j^{(N)}/s_N$ is uniformly integrable. We will prove more. Since

$$\begin{aligned}
P\big[\sigma_j^{(N)}/s_N \le 1\big] &= P\Big[\big|\xi_{x_N}^{(N)}(\Lambda(N))\big| \le 1\Big] \\
&\to \sum_{k=1}^{j} q_{\infty,k}\big(e^{-2/G}\big)
\end{aligned}$$

(a positive number) as $N \to \infty$, there exists $\delta_j > 0$ such that for all $N = 2, 4, \dots$ we have

$$P\big[\sigma_j^{(N)}/s_N \le 1\big] \ge \delta_j .$$

Now for any $A \subset \Lambda(N)$, since $\xi_t^{(N)}(A) \subset \xi_t^{(N)}(\Lambda(N))$, we must also have for all $N = 2, 4, \dots,$

$$P\big[\sigma_j^{(N)}/s_N \le 1 \big| \xi_0^{(N)} = A\big] \ge \delta_j .$$

This inequality, the Markov property and iteration lead to

$$P\big[\sigma_j^{(N)}/s_N \ge n\big] \le \big(1 - \delta_j\big)^n, \tag{5.2}$$

which certainly implies uniform integrability. This finishes the proof of Theorem 6, $d \ge 2$. □

6. Duality and the proofs of Theorems 1–3, $d \geq 2$. Duality is perhaps the chief tool used in analyzing the voter model. It is well documented in the literature, so we will only state the results we need and refer the reader to Griffeath (1979) and Liggett (1985) for proofs. It is convenient to write $A \subset \eta$ for sets $A \subset \Lambda$ and configurations $\eta \in \{0,1\}^{\Lambda}$ to mean $A \subset \{x\colon \eta(x) = 1\}$. The first duality equation we need is

$$P_{\eta}[B \subset \eta_t] = P[\xi_t(B) \subset \eta], \tag{6.1}$$

where P_{η} indicates that the voter model η_t starts with $\eta_0 = \eta$. If we start η_t in product measure with density θ, then summation in (6.1) leads to the second duality equation

$$P[B \subset \eta_t] = E[\theta^{|\xi_t(B)|}]. \tag{6.2}$$

If $0 \leq t_1 \leq \cdots \leq t_k$, then

$$P_{\eta}\big[B_i \subset \eta_{t_i}, 1 \leq i \leq k\big] = P\big[\xi_{t_k}(B_k, 0; B_{k-1}, t_k - t_{k-1}; \cdots; B_1, t_k - t_1) \subset \eta\big],$$

where $\xi_s(A_1, s_1; \cdots; A_k, s_k)$ is the coalescing random walk system defined in the previous section, with particles starting at the points of A_i at time s_i ("frozen" there until that time). Finally, if η_t starts in product measure with density θ, then

$$P\big[B_i \subset \eta_{t_i}, 1 \leq i \leq k\big] = E\big[\theta^{|\xi_{t_k}(B_k, 0;\, B_{k-1},\, t_k - t_{k-1};\, \cdots;\, B_1,\, t_k - t_1)|}\big]. \tag{6.3}$$

Before beginning the proofs of Theorems 1–3 we point out two related duality equations, discussed in Tavaré (1984), Cox and Griffeath (1986) and Cox and Griffeath (1988). These equations connect the Wright–Fisher diffusion Y_t and the death process D_t defined in Section 1. The equations are

$$E_{\theta}[Y_t^m] = E_m[\theta^{D_t}] = \sum_{j=1}^{m} \theta^j q_{m,j}(t), \tag{6.4}$$

$$E_{\theta}\left[\prod_{i=1}^{k} Y_{t_i}^{m_i}\right] = \sum_{j \geq 1} \theta^j q_{n_k, n_{k-1}, \ldots, n_1; j}(0, t_k - t_{k-1}, \ldots, t_k - t_1; t_k). \tag{6.5}$$

With these equations in hand and the coalescing random walk results of the previous sections we can now begin the voter model proofs.

Proof of Theorem 1. Fix $t \geq 0$. Then

$$\begin{aligned} P\big[\tau^{(N)} \leq ts_N\big] &= P\big[\eta_{ts_N}^{(N)} \equiv 1 \text{ or } 0\big] \\ &= E\big[\theta^{|\xi_{ts_N}^{(N)}(\Lambda(N))|}\big] + E\big[(1-\theta)^{|\xi_{ts_N}^{(N)}(\Lambda(N))|}\big] \\ &\to \sum_{k=1}^{\infty} \big[\theta^k + (1-\theta)^k\big] q_{\infty,k}(e^{-2t/G}) \end{aligned}$$

by using (6.2) and Theorem 6. To obtain moment convergence we note that $\tau^{(N)}$ is stochastically smaller than $\sigma_1^{(N)}$ and since $\sigma_1^{(N)}/s_N$ is uniformly integrable [recall (5.2)], so is $\tau^{(N)}/s_N$. This is enough to guarantee convergence of expectations. To see why $\tau^{(N)}$ is stochastically smaller than $\sigma^{(N)}$ we use duality and compute

$$\begin{aligned} P[\tau^{(N)} > t] &= 1 - E\left[\theta^{|\xi_t^{(N)}(\Lambda(N))|}\right] + E\left[(1-\theta)^{|\xi_t^{(N)}(\Lambda(N))|}\right] \\ &= \sum_{k=2}^{\infty} \left[\theta^k + (1-\theta)^k\right] P\left[\left|\xi_t^{(N)}(\Lambda(N))\right| = k\right] \\ &\le P\left[\left|\xi_t^{(N)}(\Lambda(N))\right| \ge 2\right] \\ &= P[\sigma^{(N)} > t]. \end{aligned}$$

Here is the computation for $E[\tau] = -G[\theta \log \theta + (1-\theta)\log(1-\theta)]$:

$$\begin{aligned} \int_0^{\infty} P[\tau > t]\, dt &= \sum_{k=2}^{\infty} \left[\theta^k + (1-\theta)^k\right] \int_0^{\infty} P_{\infty}[D_{2t/G} = k]\, dt \\ &= \sum_{k=2}^{\infty} \left[\theta^k + (1-\theta)^k\right] \frac{G}{2} E[\text{holding time in state } k] \\ &= \sum_{k=2}^{\infty} \left[\theta^k + (1-\theta)^k\right] \frac{G}{2} \binom{k}{2}^{-1} \\ &= -G[\theta \log \theta + (1-\theta)\log(1-\theta)]. \end{aligned}$$ □

Proof of Theorem 2. We will prove Theorem 2 in three steps, starting with:

Weak convergence of marginals. Fix $t \ge 0$ and $m \ge 1$. We will obtain

$$\Delta_{ts_N}^{(N)} \Rightarrow Y_{2t/G}$$

by showing that

$$E\left[\left(\Delta_{ts_N}^{(N)}\right)^m\right] \to E_\theta\left[Y_{2t/G}^m\right]. \tag{6.6}$$

To do this we choose a_N as in Theorem 4 and compute

$$\begin{aligned} E\left[\left(\Delta_{ts_N}^{(N)}\right)^m\right] &= N^{-md} \sum_{x_1,\dots,x_m \in \Lambda(N)} P\left[\eta_{ts_N}^{(N)}(x_i) = 1, 1 \le i \le m\right] \\ &= N^{-md} \sum_{\substack{x_1,\dots,x_m \in \Lambda(N) \\ |x_\alpha - x_\beta| \ge a_N,\, \alpha \ne \beta}} P\left[\eta_{ts_N}^{(N)}(x_i) = 1, 1 \le i \le m\right] + \varepsilon_N \\ &= N^{-md} \sum_{\substack{x_1,\dots,x_m \in \Lambda(N) \\ |x_\alpha - x_\beta| \ge a_N,\, \alpha \ne \beta}} E\left[\theta^{|\xi_{ts_N}^{(N)}(\{x_1,\dots,x_m\})|}\right] + \varepsilon_N \\ &= N^{-md} \sum_{\substack{x_1,\dots,x_m \in \Lambda(N) \\ |x_\alpha - x_\beta| \ge a_N,\, \alpha \ne \beta}} \sum_{j=1}^{m} \theta^j q_{m,j}(e^{-2t/G}) + \varepsilon_N \quad \text{(by Theorem 5)} \\ &\to \sum_{j=1}^{m} \theta^j q_{m,j}(e^{-2t/G}) \\ &= E_\theta\left[Y_{2t/G}^m\right] \end{aligned}$$

by (6.4), proving (6.6).

Weak convergence of finite-dimensional distributions. Fix $k \geq 2$, $m_i \geq 1$ and $0 \leq t_1 < \cdots < t_k$. We will prove

$$E\left[\left(\Delta_{t_1 s_N}^{(N)}\right)^{m_1} \cdots \left(\Delta_{t_k s_N}^{(N)}\right)^{m_k}\right] \to E_\theta\left[Y_{2t_1/G}^{m_1} \cdots Y_{2t_k/G}^{m_k}\right], \tag{6.7}$$

which is enough to prove

$$\left(\Delta_{t_1 s_N}^{(N)}, \dots, \Delta_{t_k s_N}^{(N)}\right) \Rightarrow \left(Y_{2t_1/G}, \dots, Y_{2t_k/G}\right).$$

We compute as before,

$$\begin{aligned} &E\left[\left(\Delta_{t_1 s_N}^{(N)}\right)^{m_1} \cdots \left(\Delta_{t_k s_N}^{(N)}\right)^{m_k}\right] \\ &\quad = N^{-d(m_1 + \cdots m_k)} \sum_{\substack{x^1, \dots, x^k \\ x^i = (x_1^i, \dots, x_m^i),\ x_j^i \in \Lambda(N)}} P\left[\eta_{t_i s_N}^{(N)}\left(x_j^i\right) = 1, 1 \leq i \leq k, 1 \leq j \leq m_i\right] \\ &\quad = \Sigma^1 + \Sigma^2, \end{aligned}$$

where Σ^1 contains all the $x^1, \dots, x^k$ such that $|x_\alpha^i - x_\beta^i| \geq a_N$, $\alpha \neq \beta$, and Σ^2 contains all the other terms. By counting it is easy to see that $\Sigma^2 = \varepsilon_N$ and by the duality equation (6.3), with $B^i = \{x_j^i, 1 \leq j \leq m_i\}$, a typical term in Σ^1 is

$$\begin{aligned} &P\left[\eta_{t_i s_N}^{(N)}\left(x_j^i\right) = 1, 1 \leq i \leq k, 1 \leq j \leq m_i\right] \\ &\quad = E\left[\theta^{|\xi_{t_k s_N}^{(N)}(B^k, 0;\ B^{k-1}, t_k - t_{k-1};\ \cdots;\ B^1, t_k - t_1)|}\right] \\ &\quad \to \sum_m \theta^m q_{n_k, n_{k-1}, \dots, n_1; m}\left(0, 2(t_k - t_{k-1})/G, \dots, 2(t_k - t_1)/G; t_k\right) \\ &\quad = E_\theta\left[Y_{2t_1/G}^{m_1} \cdots Y_{2t_k/G}^{m_1}\right] \end{aligned}$$

by (3.11) and (6.5).

Tightness. It is possible to obtain tightness on path space using Corollary 8.7 of Ethier and Kurtz (1986), Chapter 4 by showing that if f: $[0,1] \to \mathbb{R}$ is continuous, then for each $t \geq 0$,

$$\lim_{N \to \infty} \sup_\eta \left| E_\eta\left[f\left(\Delta_{t s_N}^{(N)}\right)\right] - E_{|\eta|/N^d}\left[f\left(Y_{2t/G}\right)\right]\right| = 0, \tag{6.8}$$

where $\eta \in \{0,1\}^{\Lambda(N)}$, $|\eta| = \Sigma_x \eta(x)$ and the subscripts indicate that we are starting the voter model with $\eta_0^{(N)} = \eta$ and the Wright–Fisher diffusion with $Y_0 = |\eta|/N^d$. Since each continuous f is the uniform limit of a sequence of polynomials, it suffices to show that for each $n \geq 1$, (6.8) holds with $f(x) = x^n$. But D. Aldous (personal communication) has pointed out that this is not necessary. Since $M^N(t) = \Delta_{t s_N}^{(N)}$ is a *martingale* for each N, $M^N(t)$ is *bounded* in N for each t and Y_t is *continuous*, tightness and (1.7) follow from convergence of the finite-dimensional distributions (already proved) and Proposition 1.2 of

Aldous (1989). It is easy to see that $M^N(t)$ is a martingale for fixed N, since the dynamics imply that if $\eta_t^{(N)} = \eta$ and

$$r(\eta) = \sum_{x,\, y \in \Lambda(N)} \eta(x)(1-\eta(y))p^{(N)}(x,y),$$

then $|\eta_t^{(N)}| \to |\eta_t^{(N)}| + 1$ at rate $r(\eta)$ *and* $|\eta_t^{(N)}| \to |\eta_t^{(N)}| - 1$ at rate $r(\eta)$. □

PROOF OF THEOREM 3. By inclusion–exclusion it suffices to show that

$$(6.9)\quad P\left[\eta_{t_N}^{(N)}(x) = 1,\, x \in A\right] \to \int_{[0,1]} P\left[Y_{2t/G} \in d\theta'\right]\nu_{\theta'}[\eta(x) = 1,\, x \in A].$$

We will assume that $t_N/N^d \to t \in (0,\infty)$; the two remaining cases are easy to handle. In order to prove (6.9) we need a characterization of the ν_θ of Theorem 0.

Let $\xi_\infty(A) = \lim_{t\to\infty}|\xi_t(A)|$ and let $p_n(A) = P[\xi_\infty(A) = n]$. By Theorem 0 and the duality equation (6.2),

$$\begin{aligned}\nu_\theta[\eta(x) = 1,\, x \in A] &= \lim_{t\to\infty} E\left[\theta^{|\xi_t(A)|}\right]\\ &= \sum_{n=1}^{|A|} p_n(A)\theta^n.\end{aligned}$$

The duality equation (6.4) now implies that the right-hand side of (6.11) equals

$$\sum_{n=1}^{|A|} p_n(A)E_n\left[\theta^{D(2t/G)}\right].$$

This and an application of duality to the left-hand side of (6.9) show that it suffices to prove

$$(6.10)\qquad E\left[\theta^{|\xi_{t_N}^{(N)}(A)|}\right] \to \sum_{n=1}^{|A|} p_n(A)E_n\left[\theta^{D(2t/G)}\right].$$

To do this we introduce a collection of independent random walks on $\mathbb{Z}^d$, $\{X_t(x),\, x \in \mathbb{Z}^d\}$, where $X_0(x) = x$. Standard random walk calculations can be used to show that we can find $T_N \to \infty$, $T_N = o(N^d)$ and $a_N \to \infty$, $a_N = o(N)$ as $N \to \infty$ such that for all $x \in A$,

$$(6.11)\qquad P\left[|X_t(x)| \le \sqrt{N}/\log N,\, 0 \le t \le T_N\right] \to 1,$$

$$(6.12)\qquad P\left[|X_{T_N}(x)| \le a_N\right] \to 0.$$

Now define random walks $X_t^{(N)}(x)$ on $\Lambda(N)$ by

$$X_t^{(N)}(x) = \left(X_t(x) \bmod N\right) - N/2.$$

Using the $X_t(x)$ and $X_t^{(N)}(x)$ it is clear that we can construct the processes $\xi_t(A)$ and $\xi_t^{(N)}(A)$ such that $\xi_t(A) = \xi_t^{(N)}(A)$ for all $t \le \gamma_N$, where $\gamma_N = \inf\{t \ge 0\colon |\xi_t(x)| \ge \sqrt{N}/\log N$ for some $x \in A\}$.

Letting a tilde ($\tilde{\ }$) indicate a summation over $B = \{x_1, \dots, x_n\}$ such that $|x_i - x_j| \geq a_N$ for $i \neq j$, we have

$$E\left[\theta^{|\xi_{t_N}^{(N)}(A)|}\right] = \sum_{n=1}^{|A|} \sum_{|B|=n} P\left[\xi_{T_N}^{(N)}(A) = B\right] E\left[\theta^{|\xi_{t_N - T_N}^{(N)}(B)|}\right] \quad \text{(Markov property)}$$

$$= \sum_{n=1}^{|A|} \sum_{|B|=n} P\left[\xi_{T_N}(A) = B\right] E\left[\theta^{|\xi_{t_N - T_N}^{(N)}(B)|}\right] + \varepsilon_N \quad \text{[by (6.9)]}$$

$$= \sum_{n=1}^{|A|} \tilde{\sum_{|B|=n}} P\left[\xi_{T_N}(A) = B\right] E\left[\theta^{|\xi_{t_N - T_N}^{(N)}(B)|}\right] + \varepsilon_N \quad \text{[by (6.12)]}$$

$$= \sum_{n=1}^{|A|} P\left[\left|\xi_{T_N}(A)\right| = n\right] E_n\left[\theta^{D(2t_N/GN^d - 2T_N/GN^d)}\right] + \varepsilon_N$$

(by Theorem 5)

$$\to \sum_{n=1}^{|A|} p_N(A) E_n\left[\theta^{D(2t/G)}\right]$$

as required. □

7. $d = 1$. The techniques we have used in the previous sections to analyze the voter model and coalescing random walk for $d \geq 2$ are not appropriate for $d = 1$. It is not a technical failure, but rather the behavior of our processes which differs substantially in these two regimes. Fortunately there are methods developed for the $d = 1$ case by others [Bramson and Griffeath (1980a) and Arratia (1979)] which we can adapt to the problems considered here. In addition, there is a beautiful observation of D. Aldous (personal communication) that leads to the explicit formulas for σ_1 in Theorem 6.

Arratia (1979) constructs a system c_t of coalescing Brownian motions on the line $\mathbb{R}$. Particles execute independent Brownian motions until they meet, at which time they coalesce into a single Brownian motion. The remarkable feature of this process is that the system starts at time 0 with a particle located at every $x \in \mathbb{R}$, and by each positive time t the system has only finitely many particles in every bounded set. Furthermore, Arratia proves an invariance principle for c_t and π_t, which is the point process determined by the particles in c_t.

Appropriate modifications of Arratia's work can be used to define a system $c_t = \{c_t(x), x \in (-\frac{1}{2}, \frac{1}{2}]\}$ of coalescing Brownian motions on the interval $(-\frac{1}{2}, \frac{1}{2}]$ viewed as a circle of circumference 1, where $c_t(x)$ is the position at time t of the Brownian motion started at x. The process can be visualized as a system of cannibalistic ants crawling down a tin can. Letting $\pi_t = \{c_t(x)\colon x \in [-\frac{1}{2}, \frac{1}{2})\}$, Arratia's work implies that as $N \to \infty$,

$$N^{-1}\xi_{tN^2}^{(N)}(\Lambda(N)) \Rightarrow \pi_t \tag{7.1}$$

as processes.

PROOF OF THEOREM 1, $d = 1$. Using duality and (7.1) we compute

$$\begin{aligned} P[\tau^{(N)}/N^2 \le t] &= P\left[\eta^{(N)}_{tN^2} \equiv 1 \text{ or } 0\right] \\ &= E\left[\theta^{|\xi^{(N)}_{tN^2}(\Lambda(N))|}\right] + E\left[(1-\theta)^{|\xi^{(N)}_{tN^2}(\Lambda(N))|}\right] \\ &\to \sum_{k=1}^{\infty} \left[\theta^k + (1-\theta)^k\right] P[|c_t| = k], \end{aligned}$$

thus $\tau^{(N)}/N^2 \Rightarrow \tau$. As in Section 5, $\tau^{(N)}$ is stochastically smaller than $\sigma_1^{(N)}$ and $\sigma_1^{(N)}/N^2$ is uniformly integrable. Thus $E[\tau^{(N)}/N^2] \to E[\tau]$. □

Similar remarks prove the one-dimensional version of Theorem 6. The existence of the limit of $\sigma_1^{(N)}/N^2$ can be demonstrated without using Arratia's work; the key fact is Aldous' observation that for each N,

$$\sigma_1^{(N)} =_{\mathscr{D}} \tfrac{1}{2}\bar{\sigma}^{(N)}, \tag{7.2}$$

where $\bar{\sigma}^{(N)}$ is the time it takes ($d = 1$) simple random walk starting at 1 to reach N *conditioned* on reaching N before 0. We will explore the consequences of (7.2) now, leaving its proof for the next section.

It is a simple matter [follow Problem 6 of Itô and McKean (1974), page 29] to compute the distribution of $\bar{\sigma}^{(N)}$, at least in terms of Laplace transforms. One obtains

$$E\left[e^{-\alpha\bar{\sigma}^{(N)}}\right] = \frac{1}{N}\frac{\phi^{-1}(\alpha) - \phi^{1}(\alpha)}{\phi^{-N}(\alpha) - \phi^{N}(\alpha)}, \qquad \alpha > 0, \tag{7.3}$$

where $\phi(\alpha) = 1 + \alpha - \sqrt{2\alpha + \alpha^2}$. Using (7.3) it is straightforward to check that as $N \to \infty$,

$$E\left[e^{-\alpha\bar{\sigma}^{(N)}/N^2}\right] \to \frac{\sqrt{2\alpha}}{\sinh\sqrt{2\alpha}}, \qquad \alpha > 0$$

and

$$E\left[\bar{\sigma}^{(N)}/N^2\right] \to \tfrac{1}{3},$$

finishing the proof of Theorem 6.

8. Multitype voter models. As in Cox and Griffeath (1988) one can consider the multitype voter model or *stepping stone model.* Given $\Lambda \subset \mathbb{Z}^d$ and $\kappa < \infty$ the κ-type voter model η_t on $\Lambda \subset \mathbb{Z}^d$ with transition matrix p^Λ has state space $\{0, 1, \dots, \kappa - 1\}^\Lambda$, and makes transitions

$$\eta_t^\Lambda(x) \to i \quad \text{at rate } \sum_{y \in \Lambda} p^\Lambda(x, y) 1\left(\{\eta_t^\Lambda(y) = i\}\right) \tag{8.1}$$

for $i \neq \eta_t^\Lambda(x)$. As before, we use η_t to denote the process when $\Lambda = \mathbb{Z}^d$ and $\eta_t^{(N)}$ when $\Lambda = \Lambda(N)$, governed by the transition function of simple symmetric random walk. For $\theta = (\theta_1, \dots, \theta_{\kappa-1})$ let μ_θ denote product measure on $\{0, 1, \dots, \kappa - 1\}^\Lambda$, $\mu_\theta(\eta(x) = i) = \theta_i$.

We may also consider the case $\kappa = \infty$, in which case it is natural to take $(\mathbb{Z}^d)^{\mathbb{Z}^d}$ as the state space and let the initial state satisfy $\eta(x) = x$, $x \in \mathbb{Z}^d$. The dynamics are as in (8.1), except i is now a point of $\mathbb{Z}^d$. This model is sometimes called the stepping stone model [see Cox and Griffeath (1987) for a brief discussion of its history and a list of references].

What can one say about the behavior of large finite systems for these κ-type voter models? The answer is: essentially the same things (appropriately modified) as when $\kappa = 2$. Extensions of this type are carried out in Cox and Griffeath (1988) in studying the rate of clustering of the voter model in two dimensions. We will not give detailed proofs, as the $\kappa < \infty$ is fairly easy to handle, while the $\kappa = \infty$ case requires more details than is appropriate to include here. We will discuss only the $d \geq 2$ case. One of our main reasons in stating these extensions is to write down a few formulas that are useful for comparisons with computer simulations.

$\kappa < \infty$. The analogous versions of Theorems 1 and 2 are true. Let $\tau_j^{(N)}$ be the time it takes the process to reach exactly j types, i.e.,

$$\tau_j^{(N)} = \inf\{t \geq 0; \exists A \subset \{0,1,\dots,\kappa-1\}, |A| = j, \\ \eta_t^{(N)}(x) \in A \text{ for all } x \in \Lambda(N)\}$$

and let $\Delta_t^{(N)}$ be the κ-vector $(\Delta_t^{(N)}(0), \Delta_t^{(N)}(1), \dots, \Delta_t^{(N)}(\kappa-1))$, where

$$\Delta_t^{(N)}(i) = N^{-d} \sum_{x \in \Lambda(N)} 1(\eta_t^{(N)}(x) = i).$$

Let Y_t be the κ-type Wright–Fisher diffusion which has generator

$$\frac{1}{2} \sum_{i,j=0}^{\kappa-1} \gamma_i[\delta_{ij} - \gamma_j] \frac{\partial^2}{\partial \gamma_i \partial \gamma_j}$$

and lives on the state space $\{\gamma = (\gamma_0, \dots, \gamma_{\kappa-1}): \gamma_i \geq 0, \Sigma\gamma_i = 1\}$.

By using techniques of this paper and of Cox and Griffeath (1988), one can prove: If $\eta_t^{(N)}$ has initial distribution μ_θ, then there are random variables τ_j such that

$$\tau_j^{(N)}/s_N \Rightarrow \tau_j \quad \text{and} \quad E[\tau_j^{(N)}/s_N] \Rightarrow E[\tau_j] \tag{8.2}$$

and

$$\Delta_{ts_N}^{(N)} \Rightarrow Y_{2t/G}, \qquad Y_0 = (\theta_0, \dots, \theta_{\kappa-1}) \tag{8.3}$$

as processes. The simplest way to approach (8.2) seems to be showing that for any $A \subset \{0,1,\dots,\kappa-1\}$,

$$P[\text{for all } x \in \Lambda(N), \eta_{ts_N}^{(N)}(x) \notin A] \to \sum_{j=1}^{\infty} \Big[\sum_{i \notin A} \theta_i\Big]^j q_{\infty,j}(2s/G)$$

(which follows from Theorem 6) and using inclusion–exclusion to get an explicit representation for the distribution of τ_j. The proof of (8.3) involves an extension of the duality equations (6.4) and (6.5).

The case $j = 1$ in (8.2) is particularly simple:

$$P[\tau_1 \le s] = \sum_{\ell=1}^{\infty} \left(\sum_{i=0}^{\kappa-1} \theta_i^{\ell} \right) q_{\infty,\ell}(2s/G)$$

and consequently

$$E[\tau_1] = -G \sum_{i=0}^{\kappa-1} \theta_i \log \theta_i.$$

Assuming further that $\theta = (\kappa^{-1}, \dots, \kappa^{-1})$ we have

$$E[\tau_1] = G(\kappa - 1)\log\frac{\kappa}{\kappa - 1}. \tag{8.4}$$

It is interesting to compare the $\kappa = 4$ case of (8.2) and (8.4) with the D2VOTER4 simulation of Durrett (1987), which is a simulation of the two-dimensional voter model with $\kappa = 4$ on the torus of side $N = 25$. The simulation seems to produce sample means of $\tau_1^{(N)}/s_N$ rather close to the value in (8.3).

$\kappa = \infty$. In this case (8.2) is also true, and the distribution of τ_j is the same as that of σ_j of Theorem 6, namely

$$P\left[\tau_j \le s\right] = \sum_{k=1}^{j} q_{\infty,k}(2s/G).$$

The version of (8.3) that is true with $\kappa = \infty$ takes a little explaining. For $y \in \mathbb{Z}^d$ let

$$\Delta_t^{(N)}(y) = N^{-d} \sum_{x \in \Lambda(N)} 1\left(\eta_t^{(N)}(x) = y\right).$$

Using the notation of Ethier and Kurtz (1981) let $\rho\Delta_t^{(N)} = (\delta_1, \delta_2, \dots)$ be the $\Delta_t^{(N)}(y)$ arranged in *decreasing* order and viewed as an element of the infinite-dimensional simplex $\nabla_\infty = \{\gamma = (\gamma_1, \gamma_2, \dots)\colon\ \gamma_1 \ge \gamma_2 \ge \cdots \ge 0,\ \sum_{j=1}^{\infty}\gamma_j = 1\}$. Ethier and Kurtz (1981) study a class of diffusions which live on ∇_∞, including the one with generator

$$\frac{1}{2} \sum_{i,j=0}^{\infty} \gamma_i\left[\delta_{ij} - \gamma_j\right] \frac{\partial^2}{\partial \gamma_i \, \partial \gamma_j}$$

defined on an appropriate domain. A remarkable fact is that the diffusion Y_t with this generator can be started at $(0, 0, \dots)$, in which case it jumps instantaneously into ∇_∞. Using the techniques of this paper and of Cox and Griffeath (1988) one can prove that as $N \to \infty$,

$$\rho\Delta_{ts_N}^{(N)}(y) \Rightarrow Y_{2t/G}. \tag{8.5}$$

Finally, here is the proof of (7.2): Consider the voter model $\eta_t^{(N)}$ on $\Lambda(N)$ with all types distinct, indexed by $\Lambda(N)$ itself (the $\kappa = \infty$ case). In this setting the duality equation (6.1) becomes

$$P\left[\eta_t^{(N)}(x) = y, \text{all } x \in B\right] = P\left[\xi_t^{(N)}(B) = \{y\}\right].$$

Using this duality we can write

$$\begin{aligned} P\left[\sigma_1^{(N)} \le t\right] &= P\left[\left|\xi_t^{(N)}(\Lambda(N))\right| = 1\right] \\ &= P\left[\eta_t^{(N)} \text{ has exactly one type left}\right] \\ &= \sum_{x \in \Lambda(N)} P\left[\eta_t^{(N)} \equiv x \text{ on } \Lambda(N)\right] \\ &= NP\left[\eta_t^{(N)} \equiv 0 \text{ on } \Lambda(N)\right]. \end{aligned}$$

Now $\{x\colon\ \eta_t^{(N)}(x) = 0\}$ is always an interval and $|\{x\colon\ \eta_t^{(N)}(x) = 0\}|$ is a rate 2 random walk on $\{0, 1, \dots, N\}$ starting at 1 with absorption at 0 and N. So if $\tilde{\sigma}^{(N)}$ is the time it takes such a random walk to get absorbed, then

$$P\left[\sigma_1^{(N)} \le t\right] = NP\left[\tilde{\sigma}^{(N)} \le t, \text{absorption at } N\right],$$

which is the same as (7.2), since the walk hits N before 0 with probability $1/N$.

Acknowledgments. It is a pleasure to thank David Aldous, Maury Bramson, Rick Durrett, David Griffeath, Andreas Greven and Frank Spitzer for their interest in, and help with, this work.

REFERENCES

ALDOUS, D. (1989). Stopping times and tightness. II. *Ann. Probab.* **17** 586–595.

ARRATIA, R. (1979). Coalescing Brownian motion and the voter model on $\mathbb{Z}$. Ph.D. dissertation, Univ. Wisconsin–Madison.

BHATTACHARYA, R. N. and RAO, R. R. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

BRAMSON, M. and GRIFFEATH, D. (1979). Renormalizing the 3-dimensional voter model. *Ann. Probab.* **7** 418–432.

BRAMSON, M. and GRIFFEATH, D. (1980a). Clustering and dispersion rates for some interacting particle systems on $\mathbb{Z}^1$. *Ann. Probab.* **8** 183–213.

BRAMSON, M. and GRIFFEATH, D. (1980b). Asymptotics for interacting particle systems on $\mathbb{Z}^d$. *Z. Wahrsch. verw. Gebiete* **53** 183–196.

BRAMSON, M., COX, J. T. and GRIFFEATH, D. (1986). Consolidation rates for two interacting systems in the plane. *Probab. Theory Related Fields* **73** 613–625.

CASSANDRO, M., GALVES, A., OLIVIERI, E. and VARES, M. E. (1984). Metastable behavior of stochastic dynamics: A pathwise approach. *J. Statist. Phys.* **35** 603–634.

CLIFFORD, P. and SUDBURY, A. (1973). A model for spatial conflict. *Biometrika* **60** 581–588.

COX, J. T. and GREVEN, A. (1988). On the long term behavior of some finite particle systems. Unpublished.

COX, J. T. and GRIFFEATH, D. (1983). Occupation time limit theorems for the voter model. *Ann. Probab.* **11** 876–893.

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

COX, J. T. and GRIFFEATH, D. (1987). Recent results for the stepping stone model. In *Percolation Theory and Ergodic Theory of Infinite Particle Systems* (H. Kesten, ed.). *IMA Volumes in Mathematics and Its Applications* **8** 73–83. Springer, New York.

COX, J. T. and GRIFFEATH, D. (1988). Mean field asymptotics for the planar stepping stone model. Unpublished.

DEN HOLLANDER, W. TH. F. and KASTELEYN, P. W. (1982). Random walks with "spontaneous emission" on lattices with periodically distributed imperfect traps. *Physica A* **112A** 523–543.

DOBRUSHIN, R. L. (1971). Markov processes with a large number of locally interacting components: Existence of a limit process and its ergodicity. *Problems Inform. Transmission* **7** 149–164.

DONNELLY, P. and WELSH, D. J. A. (1983). Finite particle systems and infection models. *Math. Proc. Cambridge Philos. Soc.* **94** 167–182.

DURRETT, R. (1987). *IPSMovies*. Wadsworth, Monterey, Calif.

DURRETT, R. and LIU, X.-F. (1988). The contact process on a finite set. *Ann. Probab.* **16** 1158–1173.

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

ETHIER, S. N. and KURTZ, T. G. (1981). The infinitely-many-neutral-alleles diffusion model. *Adv. in Appl. Probab.* **13** 429–454.

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

FLATTO, L., ODLYZKO, A. M. and WALES, D. B. (1985). Random shuffles and group representations. *Ann. Probab.* **13** 154–178.

GRIFFEATH, D. (1979). *Additive and Cancellative Interacting Particle Systems. Lecture Notes in Math.* **724**. Springer, New York.

GRIFFEATH, D. (1981). The basic contact process. *Stochastic Process. Appl.* **11** 151–186.

HARRIS, T. E. (1974). Contact interactions on a lattice. *Ann. Probab.* **2** 969–988.

HARRIS, T. E. (1978). Additive set-valued Markov processes. *Ann. Probab.* **4** 355–378.

HOLLEY, R. A. and LIGGETT, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3** 643–663.

ITÔ, K. and MCKEAN, H. P., JR. (1974). *Diffusion Processes and Their Sample Paths*, 2nd ed. Springer, New York.

KINGMAN, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248.

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

MONTROLL, E. W. (1969). Random walks on lattices III: Calculation of first-passage times with applications to exciton trapping on photosynthetic units. *J. Math. Phys.* **10** 753–765.

SCHONMANN, R. H. (1985). Metastability for the contact process. *J. Statist. Phys.* **41** 445–464.

SPITZER, F. L. (1976). *Principles of Random Walk*, 2nd ed. Springer, New York.

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

DEPARTMENT OF MATHEMATICS
SYRACUSE UNIVERSITY
SYRACUSE, NEW YORK 13244