

## SPATIAL MODELS FOR SPECIES AREA CURVES

BY MAURY BRAMSON,<sup>1</sup> J. THEODORE COX<sup>2</sup> AND RICHARD DURRETT<sup>3</sup>

*University of Wisconsin–Madison, Syracuse University and  
Cornell University*

The relationship between species number and area is an old problem in biology. We propose here an interacting particle system—the multitype voter model with mutation—as a mathematical model to study this problem. We analyze the species area curves of this model as the mutation rate  $\alpha$  tends to zero. We obtain two basic types of behavior depending on the size of the spatial region under consideration. If the region is a square with area  $\alpha^{-r}$ ,  $r > 1$ , then, for small  $\alpha$ , the number of species is of order  $\alpha^{1-r}(\log \alpha)^2$ , whereas if  $r < 1$ , the number of species is bounded.

**1. Introduction.** The relationship between species number and area has puzzled naturalists since the early 1800s. Watson (1835) says “On the average, a single county appears to contain nearly one half the number of species found in Britain; and it would, perhaps, not be a very erroneous guess to say that a single mile may contain half the species of a county.” Arrhenius (1921) was the first to formalize this by saying that the logarithm of the number of species was a linear function of the logarithm of the area or, equivalently, species number  $S$  and area  $A$  are related by a power law  $S = cA^z$ . While some biologists have suggested other curves, the Arrhenius relationship is the most accepted. See Connor and McCoy (1979) for a summary and analysis of more than 100 studies. A more recent survey can be found in Williamson (1988).

Although there is general agreement about the form of the species area curve, there has been much debate over the last 50 years about the causes of this relationship. For instance, Preston (1962) argued that if species abundances follow a lognormal distribution and each area is a random sample from the larger population, then  $S = cA^z$  with  $z \approx 0.26$ . MacArthur and Wilson (1967) proposed in their study of “island biogeography,” that on each island, the number of species represented an equilibrium between immigration and extinction. Hubbell (1992) was the first to suggest that one could derive species area relationships from a spatially explicit model. Unaware of the interacting particle systems literature, he constructed a fairly compli-

---

Received April 1995; revised January 1996.

<sup>1</sup>Supported in part by NSF Grant DMS-93-00612.

<sup>2</sup>Supported in part by NSF Grants DMS-93-03233 and BIR-93-21807.

<sup>3</sup>Supported in part by NSF Grant DMS-93-01070 and the Army Research Office through the Mathematical Sciences Institute at Cornell.

AMS 1991 *subject classifications*. 60K35, 92D25.

*Key words and phrases*. Species area curves, multitype contact process, multitype voter model, coalescing random walk.

cated discrete time model that allowed several species to inhabit each site, with the probability of local extinction per species being inversely proportional to the number of species at the site. He then studied the system using simulation.

In this paper, we will introduce a more elementary interacting particle system model and study its species area curves. We represent space by the two-dimensional integer lattice  $\mathbf{Z}^2$ , imagining our terrain to be divided into square cells, each of which can be occupied by at most one species. We are interested in systems where species compete and occasionally mutate into new types. In this context, it would be natural to formulate a contact process type model in which individuals can die, give birth onto vacant sites and mutate. However, to obtain a computationally tractable model, we will simplify further and formulate a voter type model in which each site is always occupied, but species can displace others in nearby sites. Results of Neuhauser (1992) suggest this change does not affect the qualitative properties of the species area curves. However, the proofs we present here would become considerably more complicated. We restrict ourselves in this paper to mathematical aspects of the problem, leaving the biological implications to be discussed in the companion paper by Durrett and Levin (1996).

In the *multitype voter model with mutation*, the state of the system at time  $t$  is given by a random function  $\xi_t: \mathbf{Z}^2 \rightarrow (0, 1)$ , with  $\xi_t(x)$  being the type of the individual at  $x$  at time  $t$ . We index our species by values  $w$  in the interval  $(0, 1)$ , so we can pick new species at random from the set of possibilities without duplicating an existing species. To formulate the dynamics, we first introduce the dispersal function  $p$ , where  $p(y - x)$  gives the probability that an offspring born at  $x$  is sent to  $y$ . For simplicity, we will make the natural assumption that  $p$  only depends upon distance. That is,  $p(y - x) = \rho(|y - x|)$ . The two mechanisms in the model—birth (with displacement) and mutation—are described by the following rules.

1. Each site  $y$  is invaded at rate 1, changing its value to the state of a site  $x$ , chosen with probability  $p(y - x)$ .
2. Each site  $y$  mutates at rate  $\alpha$ , changing to a new type  $w'$ , chosen uniformly on  $(0, 1)$ .

The voter model was first studied by Clifford and Sudbury (1973) and Holley and Liggett (1975). It and its generalizations treated in the genetics literature [see, e.g., the papers of Sawyer cited in Liggett (1985)] are mathematically tractable because there is a *duality* that allows us to reformulate questions in terms of random walks. In Section 2, we will define the voter model from a *percolation substructure* and construct its dual process, a *coalescing random walk system*  $\zeta_s^{A,t}$ ,  $0 \leq s \leq t$ , defined for each  $A \subset \mathbf{Z}^d$ . Such a system starts from  $\zeta_0^{A,t} = A$  and consists of particles which execute rate-1 independent random walks with jump kernel  $p(x)$ , except that particles coalesce when one particle jumps onto a site occupied by another particle. Since the finite-dimensional distributions of  $\zeta_s^{B,t}$  for  $s \leq t$  do not depend on  $t$ , we can let  $\zeta_s^B$  denote a process defined for all  $s \geq 0$  with these finite-

dimensional distributions. The percolation substructure also has *mutation* event markers, which occur independently at each site at rate  $\alpha$  and have independent uniform random variables attached to them. The algorithm for determining  $\xi_t(x)$  is the following. If a mutation event marker occurs anywhere on the path  $(\zeta_s^{(x),t}, 0 \leq s \leq t)$ , then the state of  $\xi_t(x)$  is the value of the uniform random variable attached to the first marker encountered. Otherwise,  $\xi_t(x) = \xi_0(y)$ , where  $y = \zeta_t^{(x),t}$ . The first application of this construction in Section 2 proves the following proposition.

PROPOSITION 1. *The multitype voter model with mutation has a unique stationary state  $\xi_\infty$ . Furthermore, for any initial  $\xi_0, \xi_t \Rightarrow \xi_\infty$  as  $t \rightarrow \infty$ .*

We use  $\Rightarrow$  to denote weak convergence, which in this setting is just convergence of finite-dimensional distributions, and we refer to  $\xi_\infty$  as an equilibrium.

The model introduced above ignores habitat diversity, but allows us to consider the question: what would species area curves look like in a homogeneous world where the only forces at work are mutation, colonization and extinction? Since we expect the mutation rate  $\alpha$  to be small, we will look at the behavior of the species area curves for the equilibria  $\xi_\infty$  as  $\alpha \rightarrow 0$  and space is rescaled in a suitable way.

The basic result we need for coalescing random walks is the asymptotic rate at which the density of particles goes to 0. This rate, obtained by Bramson and Griffeath (1980), is used in the proofs of Theorems 1 and 2. In the proof of Theorem 3, we also make use of some results of Bramson, Cox and Griffeath (1986). Both papers treat only the case where the kernel  $p(x)$  is nearest neighbor, although the results can presumably be extended to the case that  $p(x)$  has mean zero and finite variance by repeating the existing arguments with minor modifications. However, since we do not want to do this here, we will restrict our attention to the nearest neighbor case.

We now state our results for the multitype voter model with mutation on  $\mathbf{Z}^2$ . Let  $L = 1/\sqrt{\alpha}$ , and let  $N_r$  be the number of different types (or species) in  $(-L/2, L/2]^2$  in the equilibrium  $\xi_\infty$ . Write  $a^+$  for  $\max\{a, 0\}$ .

THEOREM 1. *As  $\alpha \rightarrow 0$ ,*

$$(1.1) \quad \frac{\log N_r}{2 \log L} \rightarrow (r - 1)^+ \quad \text{in probability.}$$

In the Arrhenius relation  $S = cA^z$ , the exponent  $z$  is the slope of the line obtained from a log-log plot of species number versus area. Theorem 1 says that for  $\alpha$  close to zero, plotting  $\log(N_r)/2 \log L$  as a function of  $r$  gives a line segment on  $0 \leq r \leq 1$  with slope zero and another line segment on  $r \geq 1$  with slope 1.

It is easy to see why the curve should have slope 1 for  $r \geq 1$ . First, note that the distribution of  $\xi_\infty(x), x \in B$  ( $B$  finite), can be computed by running the dual process  $\zeta_s^B$  until all particles encounter mutations. The type at  $x$  in

equilibrium will be determined by the first mutation event that occurs along the path  $\zeta_s^{(x)}$ . The time of this event has an exponential distribution with mean  $1/\alpha$ . Since  $\zeta_s^{(x)}$  is a random walk, it will have moved a distance about  $1/\sqrt{\alpha} = L$  in that time and, with high probability, will not have moved further than  $CL$  if  $C$  is a large constant. This leads to the conclusion that states of sites in  $\xi_\infty$  which are separated by a large multiple of  $L$  are typically different. So, for length scales larger than  $L$ , the number of species should be proportional to area.

To see why the limit in Theorem 1 should be 0 for  $r = 1$  (and hence for  $r < 1$ ), we need a result from Bramson and Griffeath (1980). Let  $\bar{\zeta}_t, 0 \leq t \leq \infty$ , be a coalescing random walk system starting from  $\bar{\zeta}_0 = \mathbf{Z}^2$ . By Theorem 1' in Bramson and Griffeath (1980), for large  $t$ ,

$$E \left| \bar{\zeta}_t \cap \left( -\frac{L}{2}, \frac{L}{2} \right] \right|^2 \approx L^2 \frac{\log t}{\pi t},$$

where  $\approx$  denotes ‘‘approximate equality.’’ If  $t/L^2$  is small and  $t$  is large, then one expects that

$$|\zeta_t^{B(L)}| \approx \left| \bar{\zeta}_t \cap \left( -\frac{L}{2}, \frac{L}{2} \right] \right|^2 \approx L^2 \frac{\log t}{\pi t},$$

where  $B(L) = (-L/2, L/2]^2$ . By duality,  $N_1$  is at most the number of mutations that occur in  $\zeta_t^{B(L)}, 0 \leq t < \infty$ . (The inequality arises from the depletion of the population by mutations.) Mutations occur at rate  $L^{-2}$ . So, ignoring the effect of repeated mutations, the number of mutations that occur between some fixed large time  $K$  and time  $L^2$  should have a Poisson distribution with mean approximately equal to

$$\int_K^{L^2} \frac{\log t}{\pi t} dt \approx \frac{2}{\pi} (\log L)^2.$$

It is easy to see that the expected number of mutations by time  $K$  is bounded. The number of mutations that occur after time  $L^2$  is bounded above by the number of particles in the coalescing random walk at time  $L^2$ , which is  $O(\log L)$ . This leads to the conclusion  $N_1 = O((\log L)^2)$  and hence  $\log N_1 / \log L \rightarrow 0$  as  $L \rightarrow \infty$ . A rigorous proof is given in Section 3, which shows the above computation is essentially correct.

By sharpening our argument we can considerably improve Theorem 1. Our results for the two regimes  $r \geq 1$  and  $r < 1$  are as follows.

**THEOREM 2.** *If  $r \geq 1$ , then as  $\alpha \rightarrow 0$ ,*

$$(1.2) \quad \frac{N_r}{L^{2r-2}(\log L)^2} \rightarrow \frac{2}{\pi} \text{ in probability.}$$

**THEOREM 3.** *If  $0 \leq r < 1$ , then  $N_r$  converges in distribution as  $\alpha \rightarrow 0$ . Moreover, there are finite positive constants  $B_1$  and  $B_2$  such that for all  $0 \leq r < 1$ ,*

$$(1.3) \quad \frac{B_1}{1-r} \leq \liminf_{\alpha \rightarrow 0} EN_r \leq \limsup_{\alpha \rightarrow 0} EN_r \leq \frac{B_2}{1-r}.$$

We will give the limiting distribution of  $N_r$  in Section 5. One can also show  $\lim_{\alpha \rightarrow 0} EN_r$  exists, but to do this we would need to repeat several pages of Bramson, Cox and Griffeath (1986) with minor changes.

It is, at first sight, disappointing that our limiting curve given in Theorem 1 consists of two line segments with slopes 0 and 1, rather than a single line with an “interesting” slope. However, if we assume that the species area curve is approximately linear for  $0 \leq r \leq 1$ , then Theorem 2 can be used to arrive at a slope that depends on  $\alpha$ . See Durrett and Levin (1996) for details of this and for further discussion of the biological interpretation of our results.

In this paper we have restricted our attention to two dimensions since it is the most appropriate setting for biological applications. In the one-dimensional nearest neighbor case, the different types will occupy finite unions of intervals of length of order  $L = 1/\sqrt{\alpha}$ . So, Theorem 1 holds (without the 2 in the denominator), but in Theorem 2,  $N_1$  stays bounded, and in Theorem 3,  $N_r \rightarrow 1$  in probability. In three or more dimensions, random walks are transient, so, even when  $\alpha = 0$ , there is a nontrivial stationary distribution. Known results about the size of clusters in the voter model [see, e.g., Bramson and Griffeath (1979)] suggest that in a cube of side  $L$ , there will be on the order of  $L^{d-2}$  types. That is, on log-log paper we have a species area curve with slope  $(d - 2)/d$ .

There is a large literature in mathematical genetics concerned with models in which the “spatial” locations of individuals are taken to be exchangeable (in the usual stochastic sense). The review articles of Ewens (1990), Donnelly and Tavaré (1995) and Kreitman and Akashi (1995) are good sources for information on models of this type and contain many references. To relate our work to this literature, replace  $\mathbf{Z}^d$  with the complete graph on  $L$  sites, change the dispersal kernel to be uniform over all sites and take  $\alpha = L^{-1}$ . However, in this situation, the number of species found grows like  $\log L$  as  $L \rightarrow \infty$ ; that is, there is no power law behavior.

A second nonspatial approach to the problem is to look at the number of distinct points  $N_n = |\{X_1, \dots, X_n\}|$  in the first  $n$  members of a random sample from an infinite population. Fisher, Corbet and Williams (1943) used this as a model for the number of *Lepidoptera* species (butterflies and moths) among the first  $n$  individuals collected in a light trap. If one accepts that individuals are placed randomly in space [see Coleman (1982)], one can argue that the collectors’ curve  $N_n$  is a species area curve. There have been a number of interesting mathematical studies investigating these ideas; see Karlin (1967), Rouault (1978) and Pitman (1996). In particular, appropriate assumptions on the population distribution can lead to power law growth of  $N_n$  as  $n \rightarrow \infty$ .

It is interesting to compare and contrast the spatial and nonspatial approaches to this question. However, since our understanding of that topic depends on results on the abundance of the various species in the spatial model, we will postpone our discussion to a future paper on that topic.

The remainder of this paper is devoted to proofs. Various duality relations and the proof of Proposition 1 are given in Section 2. Theorems 1–3 are proved in Sections 3–5. Theorem 1 is, of course, a consequence of Theorem 2.

We first give the simpler result and its proof to prepare the reader for the proof of Theorem 2, which is the most technical part of the paper. It involves analyzing  $\zeta_t^{B(L)}$  over several disjoint time intervals  $[T_{i-1}, T_i]$ , with Lemma 4.3 being the most involved piece. The proof of Theorem 3 is shorter and employs some of the same arguments, together with a result from Bramson, Cox and Griffeath (1986).

**2. Duality and the proof of Proposition 1.** In this section, we first describe the construction of the voter model from a percolation substructure, and the resulting duality with coalescing random walks. Proposition 1 then follows easily. We also formulate some additional notation and results which we will use in the proofs of our theorems.

Following the approach of Griffeath (1979) or Durrett (1988), we construct the classical voter model (without mutation, only two opinions) by defining, for each site  $x$ , a rate-1 Poisson process  $\{T_n^x, n \geq 1\}$  and a collection of i.i.d. random variables  $\{Z_n^x, n \geq 1\}$  with  $P(Z_n^x = z) = p(z)$ . At time  $T_n^x$ , the voter at  $y = x + Z_n^x$  will adopt the state of the voter at  $x$ . All Poisson processes  $T_n^x$  and random variables  $Z_n^x$  are assumed to be independent. The entire collection constitutes the ‘‘percolation substructure.’’ Let  $\tilde{T}_n^y$  denote the successive times at which the voter at  $y$  adopts the state of another voter; it follows with a little work that  $\{\tilde{T}_n^y, n \geq 1\}$  are also independent rate-1 Poisson processes. To facilitate the definition of the dual, we write a  $\delta$  at  $(y, T_n^x)$  and draw an arrow from  $(x, T_n^x)$  to  $(y, T_n^x)$ . We say that there is a path up from  $(x, 0)$  to  $(y, t)$  in the percolation substructure if there is a sequence of times  $0 = s_0 < s_1 < s_2 < \dots < s_n < s_{n+1} = t$  and spatial locations  $x = x_0, x_1, \dots, x_n = y$ , so that the following statements hold:

1. For  $1 \leq i \leq n$ , there is an arrow from  $x_{i-1}$  to  $x_i$  at time  $s_i$ .
2. For  $0 \leq i \leq n$ , the vertical segments  $\{x_i\} \times (s_i, s_{i+1})$  do not contain any  $\delta$ 's.

For each set of sites  $A$ , we put  $\eta_0^A = A$  and define, for  $t > 0$ ,

$$\eta_t^A = \{y: \text{for some } x \in A \text{ there is a path up from } (x, 0) \text{ to } (y, t)\}.$$

Here  $\eta_t^A$  is the classical voter model with possible opinions 0 and 1. If  $A$  denotes the set of sites occupied by 1's at time 0, then  $\eta_t^A$  is the set of sites occupied by 1's at time  $t$ .

An important feature of this construction is that we can construct a dual process on the same percolation substructure. We reverse the directions of the arrows and define paths going down in the analogous way. For each set of sites  $B$ , for fixed  $t$  and  $0 \leq s \leq t$ , put

$$\zeta_s^{B,t} = \{x: \text{for some } y \in B, \text{ there is a path down from } (y, t) \text{ to } (x, t - s)\}.$$

Then  $\eta_t^A$  and  $\zeta_s^{B,t}$  are *dual* in the sense that for any  $0 \leq s \leq t$ ,

$$(2.1) \quad \{\eta_t^A \cap B \neq \emptyset\} = \{A \cap \zeta_t^{(B,t)} \neq \emptyset\}.$$

The finite-dimensional distributions of  $\zeta_s^{B,t}$  for  $s \leq t$  do not depend on  $t$ , so we can let  $\zeta_s^B$  denote a process defined for all  $s \geq 0$  with these finite-dimen-

sional distributions, and call  $\zeta_s^B$  the dual of  $\eta_t^A$ . It follows from (2.1) that

$$(2.2) \quad P(\eta_t^A \cap B \neq \emptyset) = P(A \cap \zeta_t^B \neq \emptyset).$$

It is easy to see that the dual process  $\zeta_s^B$  is a coalescing random walk. The individual particles in  $\zeta_s^B$  perform independent rate-1 random walks with jump distribution  $p(x)$ , with the collision rule that when two particles meet, they coalesce into a single particle.

We may also define a multitype voter model using the same percolation substructure. Assume now that the type of an individual belongs to  $(0, 1)$ , and write  $\eta_t(x)$  for the type of the individual at site  $x$  at time  $t$ . Given  $\eta_0$ , we put  $\eta_t(y) = \eta_0(x)$ , where  $x$  is the unique site such that there is a path up from  $(x, 0)$  to  $(y, t)$  in the percolation substructure.

We need to generalize our construction to include mutations. We enrich the percolation substructure with independent rate- $\alpha$  Poisson processes  $\{S_n^x, n \geq 1\}$  and i.i.d. random variables  $\{U_n^x, n \geq 1\}$  that are uniform on the interval  $(0, 1)$ . The times  $S_n^x$  are the times at which a mutation occurs at site  $x$ , and the new type is given by  $U_n^x$ . We can now define the multitype voter model with mutation  $\xi_t$  and an additional “dual” process  $\hat{\zeta}_s^{B,t}$ . Fix  $\xi_0$ , where  $\xi_0(x) \in (0, 1)$  is the type of the site  $x$  at time 0. To determine  $\xi_t(y)$ , choose the unique site  $x$  such that there is a path up from  $(x, 0)$  to  $(y, t)$ . If there is no mutation event on this path, put  $\xi_t(y) = \xi_0(x)$ . Otherwise, let  $(z, t')$  be the point on this path with the property that  $S_n^z = t'$  for some  $n$  and there are no other mutation events on the path from  $(z, t')$  up to  $(y, t)$ . Put  $\xi_t(y) = U_n^z$ . The process  $\hat{\zeta}_s^{B,t}$  is defined just as is the process  $\zeta_s^{B,t}$ , except that particles encountering mutation events disappear. Write  $\hat{\zeta}_s^B$  for the process defined for all  $s \geq 0$  with the same finite-dimensional distributions as  $\hat{\zeta}_s^{B,t}$ ,  $s \leq t$ .

PROOF OF PROPOSITION 1. We first note that if  $\hat{\zeta}_s^{(B,t)} = \emptyset$  for some  $s < t$ , then the states  $\xi_t(x)$  of the sites  $x \in B$  do not depend on the initial configuration  $\xi_0$ . The number of particles  $|\hat{\zeta}_s^{(B,t)}|$  has the same distribution as  $|\hat{\zeta}_s^B|$ . For finite  $B$ ,  $|\hat{\zeta}_s^B|$  never increases and the particles in  $\hat{\zeta}_s^B$  disappear at rate  $\alpha > 0$ , so  $P(\hat{\zeta}_s^B \neq \emptyset) \rightarrow 0$  as  $s \rightarrow \infty$ . This implies that  $\xi_t$  converges weakly as  $t \rightarrow \infty$  and the limit does not depend on the initial configuration. Since the voter model with mutation is a Feller process, the limit  $\xi_\infty$  is an equilibrium. Since the limit is independent of the initial configuration, it is unique.  $\square$

We need some additional preparatory material for the proofs of Theorems 1–3. For the process  $\xi_t$ , define

$$N_t^A = \text{the number of species (types) in set } A \text{ at time } t.$$

We will sometimes write  $N_t(A)$  for  $N_t^A$ . For  $0 \leq t_1 \leq t_2 \leq t$ , let

$$Y_{t_1, t_2}^{A,t} = \text{the number of mutations occurring on } (\zeta_s^{A,t}, t_1 \leq s < t_2),$$

$$\hat{Y}_{t_1, t_2}^{A,t} = \text{the number of mutations occurring on } (\hat{\zeta}_s^{A,t}, t_1 \leq s < t_2).$$

We will refer to  $(\hat{\zeta}_s^{A,t}, 0 \leq s \leq t)$  as the *ancestral path* and to  $(\zeta_s^{A,t}, 0 \leq s \leq t)$  as the *extended ancestral path* of  $A$ . We will make use of the shorthand

$$Y_{t_2}^{A,t} = Y_{0,t_2}^{A,t}, \quad \hat{Y}_{t_2}^{A,t} = \hat{Y}_{0,t_2}^{A,t}.$$

By inspecting the percolation substructure and the definitions of  $\xi_t$  and  $\hat{\zeta}_s^{A,t}$ , it is not difficult to see that

$$(2.3) \quad N_t^A = \hat{Y}_t^{A,t} + N_0(\hat{\zeta}_t^{A,t})$$

or, more generally,

$$(2.4) \quad N_t^A = \sum_{i=1}^I \hat{Y}_{t_{i-1},t_i}^{A,t} + N_0(\hat{\zeta}_t^{A,t})$$

for  $0 = t_0 < t_1 < \dots < t_I = t$ . Since  $\hat{Y}_{t_1,t_2}^{A,t} \leq Y_{t_1,t_2}^{A,t}$  and  $N_0(\hat{\zeta}_t^{A,t}) \leq |\zeta_t^{A,t}|$ , one has

$$(2.5) \quad N_t^A \leq \sum_{i=1}^I Y_{t_{i-1},t_i}^{A,t} + |\zeta_t^{A,t}|.$$

There is a useful “converse” to  $\hat{\zeta}_s^{A,t} \subset \zeta_s^{A,t}$ , namely,

$$(2.6) \quad |\zeta_s^{A,t}| \leq |\hat{\zeta}_s^{A,t}| + \hat{Y}_s^{A,t} \leq |\hat{\zeta}_s^{A,t}| + Y_s^{A,t}.$$

Now set

$$y_{t_1,t_2}^A = EY_{t_1,t_2}^{A,t}, \quad 0 \leq t_1 \leq t_2 \leq t, \\ z_s^A = E|\zeta_s^{A,t}|, \quad s \leq 1.$$

Taking expectations in (2.5) gives

$$(2.7) \quad EN_t^A \leq \sum_{i=1}^I y_{t_{i-1},t_i}^A + z_t^A.$$

Since mutations occur independently of  $(\zeta_s^{A,t}, 0 \leq s \leq t)$ ,

$$(2.8) \quad y_{t_{i-1},t_i}^A = \alpha \int_{t_{i-1}}^{t_i} z_s^A ds.$$

This gives, in particular,

$$(2.9) \quad EN_t^A \leq z_t^A + \alpha \int_0^t z_s^A ds.$$

Inequalities (2.5)–(2.9) will be used at various places in Sections 4 and 5.

We will also need to do some more careful estimation for  $\hat{Y}$  in Section 4. Suppose that  $|\hat{\zeta}_s^{A,t}| \geq f(s)$  holds for  $s \in [t_{i-1}, t_i]$  on some event  $G$ , for a given function  $f(s)$ . Then there is a Poisson random variable  $X$  with mean  $\alpha \int_{t_{i-1}}^{t_i} f(s) ds$  so that

$$(2.10) \quad \hat{Y}_{t_{i-1},t_i}^{A,t} \geq X \quad \text{on } G.$$

(Note that  $\hat{Y}_s^{A,t}$  is a pure birth process that increases at rate  $\alpha|\hat{\zeta}_s^{A,t}|$ , and one can compare this with a Poisson process by rescaling  $s$ .) There is an analogous upper bound when  $|\hat{\zeta}_s^{A,t}| \leq f(s)$  holds for  $s \in [t_{i-1}, t_i]$  on some event  $G$ .



We define one final auxiliary process,  $\tilde{\zeta}_s^{A,t}$ ,  $0 \leq s \leq t$ . We modify the definition of “downward paths” in the percolation substructure by allowing vertical segments  $\{x_i\} \times (s_i, s_{i+1})$  to contain a  $\delta$ , provided  $x_i \in A$ . Set  $\tilde{\zeta}_0^{A,t} = A$  and let  $\tilde{\zeta}_s^{A,t}$  denote the set of sites  $y$  such that there is a modified path down from  $A \times \{t\}$  to  $(y, t - s)$ . That is, the process  $\tilde{\zeta}_s^{A,t}$  is a coalescing random walk system such that whenever a walk leaves a site of  $A$ , another walk is created instantaneously there. (This creation of new walks always occurs at rate at most 1.) We will write  $\tilde{\zeta}_s^A$  for a process with the same finite-dimensional distributions as  $\tilde{\zeta}_s^{A,t}$ . Letting  $O$  denote the origin, it follows easily that

$$(2.11) \quad \{\eta_s^{(0)} \cap A \neq \emptyset \text{ for some } s \leq t\} = \{O \in \tilde{\zeta}_s^{A,t} \text{ for some } s \leq t\}.$$

**3. Proof of Theorem 1.**

PROOF OF THE LOWER BOUND. This part is simple. We can assume that  $r > 1$  since there is otherwise nothing to prove. Start with the process  $\xi_t$  in equilibrium and let it evolve using the percolation substructure introduced in Section 2. We call a site  $x$  good if  $S_1^x < 1 < \tilde{T}_1^x$ , that is, if the individual at  $x$  undergoes a mutation by time 1 but adopts no other individual’s state by time 1. Denote by  $G_r$  the number of good sites in  $(-L^r/2, L^r/2]^2$  and by  $N_r^1$  the number of species there for the process at time 1. Since  $N_r$  and  $N_r^1$  have the same distribution and  $G_r \leq N_r^1$ , it suffices to get a lower bound on  $G_r$ .

Different sites are good with probability  $p = e^{-1}(1 - e^{-\alpha})$  independently of one another. Thus  $G_r$  is the sum of independent Bernoulli random variables with mean  $p$  and variance  $p(1 - p)$ . So, letting  $[x]$  denote the integer part of  $x$ ,

$$(3.1) \quad EG_r = [L^r]^2 e^{-1}(1 - e^{-\alpha}) \sim e^{-1}L^{2(r-1)}, \quad \text{var}(G_r) \leq EG_r.$$

(The notation  $a_L \sim b_L$  means that  $a_L/b_L \rightarrow 1$  as  $L \rightarrow \infty$ .) Chebyshev’s inequality implies that there is a constant  $C < \infty$  so that

$$(3.2) \quad P(|G_r - EG_r| > L^{3(r-1)/2}) \leq CL^{-(r-1)} \rightarrow 0$$

as  $L \rightarrow \infty$ . Substitution of the estimate for  $EG_r$  in (3.1) into (3.2) shows that  $P(G_r < L^{2(r-1)}/3) \rightarrow 0$  as  $L \rightarrow \infty$ . This implies that  $P(N_r < L^{2(r-1)}/3) \rightarrow 0$  as  $L \rightarrow \infty$  (and hence as  $\alpha \rightarrow 0$ ) and consequently the lower bound in (1.1) holds.  $\square$

REMARK. The above argument immediately generalizes to any finite range model with mutation.

Throughout the text,  $C$  will denote a finite positive constant whose value is unimportant, and which will be allowed to change from line to line.

PROOF OF THE UPPER BOUND. Since the map  $r \rightarrow N_r$  is nondecreasing, it is clearly enough to prove the bound for  $r > 1$ . It suffices to show

$$(3.3) \quad \limsup_{L \rightarrow \infty} \frac{\log(EN_r)}{2 \log L} \leq (r - 1)^+.$$

To see this, we note that (3.3) implies that if  $\delta > 0$  and  $\alpha = 2(r-1)^+ + \delta$ , then  $EN_r \leq L^\alpha$  for large  $L$ . So Chebyshev's inequality implies  $P(N_r \geq L^{\alpha+d}) \leq L^{-\delta}$  and it follows that

$$P\left(\frac{\log N_r}{2 \log L} > (r-1)^+ + \delta\right) \rightarrow 0$$

as  $L \rightarrow \infty$  (and hence  $\alpha \rightarrow 0$ ).

To demonstrate (3.3), start  $\xi_t$  in equilibrium and let it evolve using the percolation substructure from Section 2. Here and later on we employ the notation  $B(M) = (-M/2, M/2]^2$  for  $M > 0$ . It follows from (2.9) that

$$(3.4) \quad EN_r = EN_T^{B(M)} \leq z_T^{B(M)} + \alpha \int_0^T z_s^{B(M)} ds$$

for  $T \geq 0$ , where  $M = L^r$ . Most of the work in estimating  $EN_r$  consists of obtaining the following bound on  $z_t^{B(M)}$ .

**LEMMA 3.1.** *If  $0 < \varepsilon < 1/3$  and  $M \geq 1$ , then for  $e^2 \leq t \leq M^{2(1-\varepsilon)}$  and appropriate  $C$ ,*

$$(3.5) \quad z_t^{B(M)} \leq \frac{CM^2 \log t}{t}.$$

We will first use Lemma 3.1 to prove Theorem 1 and then prove the lemma. Let  $T = L^2 = M^{2(1-\varepsilon)}$ , where  $\varepsilon = 1 - 1/r$ . (Recall that  $r > 1$ .) From (3.5), it follows that

$$(3.6) \quad z_T^{B(M)} \leq \frac{CM^2 \log(L^2)}{L^2} = 2CL^{2(r-1)} \log L.$$

The trivial bound  $z_t^{B(M)} \leq M^2 = L^{2r}$  for  $t \leq e^2$  and integration of (3.5) from  $e^2$  to  $T$  also gives

$$(3.7) \quad \alpha \int_0^T z_s^{B(M)} ds \leq \alpha CM^2 \left( e^2 + \int_{e^2}^T \frac{\log t}{t} dt \right) \leq CL^{2(r-1)} (\log L)^2,$$

with the last inequality following from the fact that  $\alpha = 1/L^2$  and  $T = L^2$ . (Recall  $C$  changes from line to line.) Plugging (3.6) and (3.7) into (3.4), we get

$$(3.8) \quad EN_r \leq CL^{2(r-1)} (\log L)^2.$$

Taking logarithms in (3.8) shows that

$$\limsup_{L \rightarrow \infty} \frac{\log EN_r}{2 \log L} \leq \lim_{L \rightarrow \infty} \frac{\log(CL^{2(r-1)} (\log L)^2)}{2 \log L} = (r-1).$$

This gives (3.3), assuming Lemma 3.1 holds.

We use the following random walk estimate in the proof of Lemma 3.1. Let  $\|x\| = \max\{|x_i|: i = 1, \dots, d\}$  and recall that  $|x| = (x_1^2 + \dots + x_d^2)^{1/2}$ .

LEMMA 3.2. *Let  $W_t$  be a finite range  $d$ -dimensional random walk that takes jumps at rate 1 and starts at  $O$ . Suppose the jump distribution satisfies  $p(x) = \rho(|x|)$  and set  $\sigma^2 = \sum_x x_i^2 p(x)$ ,  $i = 1, 2, \dots, d$ . There is a constant  $u_0$  so that if  $u \geq u_0$  and  $K \leq u^{1/4}$ , then*

$$(3.9) \quad P\left(\max_{t \leq u} \|W_t\| > K\sqrt{u}\right) \leq 4d \exp(-K^2/4\sigma^2).$$

PROOF. By considering the coordinates  $W_t^i$  separately, it is enough to prove that

$$P\left(\max_{t \leq u} W_t^1 > K\sqrt{u}\right) \leq 2 \exp(-K^2/4\sigma^2).$$

Since  $W_t^1$  is symmetric, by the reflection principle, it is enough to show

$$(3.10) \quad P(W_u^1 > K\sqrt{u}) \leq \exp(-K^2/4\sigma^2).$$

The transform of  $W_u^1$  is given by

$$E \exp(\theta W_u^1) = \exp(u(\psi(\theta) - 1)),$$

where  $\psi(\theta) = \sum_{x \in \mathbf{Z}^d} p(x) \exp(\theta x_1)$ . For  $\theta > 0$ , it follows from Chebyshev's inequality that

$$P(W_u^1 > K\sqrt{u}) \leq \exp(-\theta K\sqrt{u} + u(\psi(\theta) - 1)).$$

Now  $\psi(0) = 1$ ,  $\psi'(0) = \sum_x x_1 p(x) = 0$  and  $\psi''(0) = \sigma^2 < \infty$ , so  $\psi(\theta) - 1 \sim \sigma^2 \theta^2 / 2$  as  $\theta \rightarrow 0$ . Pick  $\theta_0 > 0$  so that  $\psi(\theta) - 1 \leq 3\sigma^2 \theta^2 / 4$  for  $0 \leq \theta \leq \theta_0$ . Taking  $\theta = K/(\sigma^2 \sqrt{u})$  and choosing  $u_0$  so that  $K/(\sigma^2 \sqrt{u_0}) \leq \theta_0$ , we have for  $u \geq u_0$ ,

$$\begin{aligned} P(W_t^1 > K\sqrt{u}) &\leq \exp\left(-\theta K\sqrt{u} + u \frac{3\sigma^2 \theta^2}{4}\right) \\ &= \exp\left(-\frac{K^2}{\sigma^2} + \frac{3K^2}{4\sigma^2}\right) = \exp\left(\frac{-K^2}{4\sigma^2}\right). \end{aligned}$$

This implies (3.10), and hence the lemma.  $\square$

Let us write  $\bar{\zeta}_t$  for  $\zeta_t^{\mathbf{Z}^2}$ . For the proof of Lemma 3.1, we need the following asymptotics from Theorem 1' of Bramson and Griffeath (1980).

$$(3.11) \quad P(O \in \bar{\zeta}_t) \sim \frac{\log t}{\pi t} \quad \text{as } t \rightarrow \infty.$$

PROOF OF LEMMA 3.1. One can decompose  $z_t^{B(M)}$  into

$$(3.12) \quad z_t^{B(M)} = E|\zeta_t^{B(M)} \cap B(2M)| + E|\zeta_t^{B(M)} \cap B(2M)^c|.$$

Since  $\zeta_t^{B(M)} \subset \bar{\zeta}_t$ , one has

$$(3.13) \quad E|\zeta_t^{B(M)} \cap B(2M)| \leq E|\bar{\zeta}_t \cap B(2M)| \leq (2M + 1)^2 P(O \in \bar{\zeta}_t).$$

On account of (3.11), this gives

$$(3.14) \quad E|\zeta_t^{B(M)} \cap B(2M)| \leq CM^2(\log t)/t$$

for  $t \geq e^2$  and appropriate  $C$ .

One can employ Lemma 3.2 for the other half of (3.12). Setting  $u = M^{2(1-\varepsilon)}$ ,  $K = M^\varepsilon/2$ ,  $d = 2$  and  $\sigma^2 = 1/2$ , it follows from Lemma 3.2 that

$$(3.15) \quad P\left(\max_{t \leq u} |\zeta_t^{(0)}| > M/2\right) \leq 8 \exp(-M^{2\varepsilon}/8).$$

(Since  $\varepsilon \leq 1/3$  is assumed,  $K \leq u^{1/4}$  holds.) There are at most  $M^2$  sites in  $B(M)$ , none of which is within distance  $M/2$  of  $B(2M)^c$ . So on account of (3.15), for  $t \leq M^{2(1-\varepsilon)}$ ,

$$(3.16) \quad \begin{aligned} E|\zeta_t^{B(M)} \cap B(2M)^c| &\leq M^2 P(\|\zeta_t^{(0)}\| > M/2) \\ &\leq 8M^2 \exp(-M^{2\varepsilon}/8). \end{aligned}$$

This gives the bound

$$(3.17) \quad E|\zeta_t^{B(M)} \cap B(2M)^c| \leq CM^2/t$$

for appropriate  $C$ . Inserting (3.14) and (3.17) into (3.12), one obtains

$$z_t^{B(M)} \leq CM^2(\log t)/t$$

This implies (3.5).  $\square$

**4. Proof of Theorem 2.** As in Section 3, we use  $\zeta_t^{B(L)}$  and  $\bar{\zeta}_t$  to denote the coalescing random walks with  $\zeta_0^{B(L)} = B(L)$  and  $\bar{\zeta}_0 = \mathbf{Z}^2$ , respectively. Intuitively, it should be the case that if  $t$  is large and  $t/L^2$  is small, then

$$|\zeta_t^{B(L)}| \approx |\bar{\zeta}_t \cap B(L)| \approx L^2 \frac{\log t}{\pi t}.$$

Much of the work done in the proof of Theorem 2 is devoted to making this precise in the form of the following two propositions.

PROPOSITION 2. For all  $\varepsilon > 0$ ,

$$(4.1) \quad P\left(\frac{\pi t |\zeta_t^{B(L)}|}{L^2 \log t} \in [(1-\varepsilon), (1+\varepsilon)] \text{ for all } t \in [\log L, L^2/(\log L)^3]\right) \rightarrow 1$$

as  $L \rightarrow \infty$

PROPOSITION 3. For all  $\varepsilon > 0$ ,

$$(4.2) \quad E|\zeta_t^{B(L)}| \leq (1 + \varepsilon)L^2(\log t)/\pi t \text{ for } t \in [\log L, L^2/(\log \log L)^2]$$

and large enough  $L$ .

We will use Propositions 2 and 3 to prove Theorem 2, and then go back and prove them.

PROOF OF THE UPPER BOUND. We employ the duality relationships in Section 2 for  $\xi_t$  started from equilibrium. Let  $T_0 = 0$ ,  $T_1 = \log K$ ,  $T_2 =$

$L^2/(\log L)^3$  and  $T_3 = L^2/(\log \log L)^2$ , where  $K = L^r$ . (Recall that  $r \geq 1$ .) Let  $Y_i = Y_{T_{i-1}, T_i}^{B(K)}$ ,  $i = 1, 2, 3$ , be the number of mutations along the extended ancestral path between times  $T_{i-1}$  and  $T_i$ . Set  $N_r^t$  equal to the number of species in  $B(L^r)$  at time  $t$ . By (2.5),

$$(4.3) \quad N_r^{T_3} \leq Y_1 + Y_2 + Y_3 + |\zeta_{T_3}^{B(K), T_3}|.$$

Since  $N_r^{T_3} =_d N_r$ , we can employ (4.3) for our bound.

We proceed to estimate the terms on the right-hand side of (4.3). As we shall see, all of the terms except  $Y_2$  are  $o(L^{2r-2}(\log L)^2)$  as  $L \rightarrow \infty$ . It is easy to see this for  $Y_1$ . Since  $|\zeta_t^{B(K)}| \leq K^2$  and mutations occur at rate  $\alpha = L^{-2}$ , we have  $EY_1 \leq rL^{2r-2} \log L$ . So,

$$(4.4) \quad Y_1/L^{2r-2}(\log L)^2 \rightarrow 0 \quad \text{in probability}$$

To analyze  $Y_2$ , we substitute  $K$  for  $L$  in Proposition 2. One gets that with probability close to 1,

$$(4.5) \quad |\zeta_t^{B(K)}| \leq (1 + \varepsilon)K^2(\log t)/\pi t \quad \text{for all } t \in [\log K, L^2/(\log L)^3].$$

[ $L^2/(\log L)^3$  is increasing in  $L$  for  $L \geq e$ .] Set

$$f(t) = (1 + \varepsilon)K^2(\log t)/\pi t.$$

The upper bound corresponding to (2.10) implies that on the event where (4.5) holds,  $Y_2$  is smaller in distribution than  $X_1$ , a Poisson random variable with mean

$$\begin{aligned} \lambda &= (1 + \varepsilon)\alpha K^2 \int_{\log K}^{L^2/(\log L)^3} \frac{\log t}{\pi t} dt \\ &= \frac{1 + \varepsilon}{2\pi} L^{2r-2} \{(2 \log L - 3 \log \log L)^2 - (\log(r \log L))^2\}. \end{aligned}$$

As  $L \rightarrow \infty$ ,

$$\pi\lambda/2(1 + \varepsilon)L^{2r-2}(\log L)^2 \rightarrow 1.$$

By Chebyshev's inequality,  $X_1/\lambda \rightarrow 1$  in probability (since  $X_1$  has mean and variance equal to  $\lambda$ ). Therefore,

$$(4.6) \quad P\left(Y_2 > \frac{2}{\pi}(1 + 2\varepsilon)L^{2r-2}(\log L)^2\right) \rightarrow 0$$

as  $L \rightarrow \infty$ .

To analyze  $Y_3$ , we substitute  $K$  for  $L$  in Proposition 3. Together with (2.8), this implies that

$$\begin{aligned} EY_3 &\leq (1 + \varepsilon)\alpha K^2 \int_{T_2}^{T_3} \frac{\log t}{\pi t} dt = \frac{1 + \varepsilon}{2\pi} L^{2r-2} \{(\log T_3)^2 - (\log T_2)^2\} \\ &= \frac{1 + \varepsilon}{2\pi} L^{2r-2} (\log T_3 + \log T_2)(\log T_3 - \log T_2) \\ &\leq \frac{1 + \varepsilon}{2\pi} L^{2r-2} (4 \log L)(3 \log \log L). \end{aligned}$$

Consequently,

$$(4.7) \quad Y_3/L^{2r-2}(\log L)^2 \rightarrow 0 \quad \text{in probability}$$

as  $L \rightarrow \infty$ . It also follows from Proposition 3, that

$$\limsup_{L \rightarrow \infty} E|\zeta_{T_3}^{B(K)}|/L^{2r-2} \log L(\log \log L)^2 \leq 2/\pi.$$

One therefore gets that

$$(4.8) \quad |\zeta_{T_3}^{B(K)}|/L^{2r-2}(\log L)^2 \rightarrow 0 \quad \text{in probability}$$

as  $L \rightarrow \infty$ . Combining (4.4), (4.6), (4.7) and (4.8), we obtain that for any  $\varepsilon > 0$ ,

$$(4.9) \quad P\left(N_r > \frac{2(1 + 3\varepsilon)}{\pi} L^{2r-2}(\log L)^2\right) \rightarrow 0 \quad \text{in probability.} \quad \square$$

PROOF OF THE LOWER BOUND. The upper bounds (4.4) and (4.6) imply that

$$(4.10) \quad Y_{T_2}^{B(K)}/L^{2r-2}(\log L)^3 \rightarrow 0 \quad \text{in probability}$$

as  $L \rightarrow \infty$ . Also, note that  $T_2 = L^2/(\log L)^3$ , so

$$(4.11) \quad \frac{L^{2r} \log T_2}{T_2} > L^{2r-2}(\log L)^3$$

for large  $L$ . Using Proposition 2 now, together with (2.6), it follows that for any  $\varepsilon > 0$ ,

$$(4.12) \quad P\left(|\hat{\zeta}_t^{B(K)}| \geq (1 - 2\varepsilon) \frac{L^{2r} \log t}{\pi t} \text{ for all } t \in [T_1, T_2]\right) \rightarrow 1$$

as  $L \rightarrow \infty$ . We denote by  $G_L$  the set on which the event in (4.12) holds for a given  $L$ .

We now mimic the reasoning leading to the upper bound (4.6). Set

$$f(t) = (1 - 2\varepsilon)L^{2r}(\log t)/\pi t.$$

Appealing to (2.10) of Section 2, we see that on  $G_L$ ,  $\hat{Y}_{T_1, T_2}^{B(K)}$  is larger in distribution than  $X_2$ , a Poisson random variable with mean

$$\begin{aligned} \lambda &= (1 - 2\varepsilon)L^{2r-2} \int_{r \log L}^{L^2/(\log L)^3} \frac{\log t}{\pi t} dt \\ &= \frac{1 - 2\varepsilon}{2\pi} L^{2r-2} \{(2 \log L - 3 \log \log L)^2 - (\log(r \log L))^2\}. \end{aligned}$$

As  $L \rightarrow \infty$ ,

$$\pi\lambda/2(1 - 2\varepsilon)L^{2r-2}(\log L)^2 \rightarrow 1$$

and  $X_2/\lambda \rightarrow 1$  in probability. So, we have

$$P\left(\hat{Y}_{T_1, T_2}^{B(K)} < \frac{2(1 - 3\varepsilon)}{\pi} L^{2r-2}(\log L)^2\right) \rightarrow 0$$

as  $L \rightarrow \infty$  for any  $\varepsilon > 0$ . Since  $N_r^{T_3} \geq \hat{Y}_{T_1, T_2}^{B(K)}$ , this implies

$$P\left(N_r < \frac{2(1 - 3\varepsilon)}{\pi} L^{2r-2} (\log L)^2\right) \rightarrow 0$$

as  $L \rightarrow \infty$ . The proof of Theorem 2 is complete, except, of course, for the proofs of Propositions 2 and 3.  $\square$

PROOFS OF PROPOSITIONS 2 AND 3. We employ several lemmas here. For Lemma 4.1, we introduce the following terminology. For given  $L > 0$ ,  $\partial B(L)$  will denote the set of points in  $\mathbf{Z}^2$  outside, but within distance 1 of  $B(L)$  (in the  $\|\cdot\|$  norm). We consider the process of coalescing random walks with creation of particles at each site of  $\partial B(L)$ ,  $\tilde{\zeta}_t^{\partial B(L)}$ , as defined in Section 2. We are interested in  $H_{L, K, u}$ , the number of random walk paths in  $\tilde{\zeta}_t^{\partial B(L)}$  which move a distance greater than  $K\sqrt{u}$  from their initial points by time  $u$ .

LEMMA 4.1. *There are constants  $u_0$  and  $C$  so that if  $u \geq u_0$ ,  $K \leq u^{1/4}$  and  $L \geq 1$ , then*

$$(4.13) \quad EH_{L, K, u} \leq CLu \exp(-K^2/2).$$

PROOF. By Lemma 3.2, the probability of a random walk moving distance greater than  $K\sqrt{u}$  by time  $u$  is at most  $8 \exp(-K^2/2)$ . For  $L \geq 1$ ,  $|\partial B(L)| \leq CL$ . Since the creation of particles at each  $x \in \partial B(L)$  occurs at rate at most 1, the number of births of such random walks on  $\partial B(L)$  over the time interval  $[0, u]$  is therefore bounded by a Poisson random variable with mean  $CLu$ . Taking expected values gives (4.13).  $\square$

Lemma 4.1 leads easily to the following bound, which will be used in the proof of Lemma 4.2.

COROLLARY 1. *Suppose that  $\underline{L} \leq L - 2\sqrt{u}K$ , where  $L$  is large,  $K \leq u^{1/4}$  and  $u \geq u_0$ . Then there is a constant  $C$  so that*

$$(4.14) \quad E|\bar{\zeta}_t \cap B(\underline{L})| - E|\zeta_t^{B(L)} \cap B(\underline{L})| \leq CLu \exp(-K^2/2) \quad \text{for all } t \leq u.$$

PROOF. It is clear from the definition of the coalescing random walk that

$$(4.15) \quad E|\bar{\zeta}_t \cap B(\underline{L})| - E|\zeta_t^{B(L)} \cap B(\underline{L})| \leq E|\zeta_t^{B(L)^c} \cap B(\underline{L})|.$$

Every random walk starting outside  $B(L)$  and entering  $B(\underline{L})$  must pass through  $\partial B(L)$ . Thus we can dominate  $\zeta_t^{B(L)^c} \cap B(\underline{L})$  by the process  $\tilde{\zeta}_t^{\partial B(L)}$  defined at the end of Section 2. Since all points in  $B(\underline{L})$  are distance greater than  $\sqrt{u}K$  away from  $B(L)^c$ ,

$$|\zeta^{B(L)^c} \cap B(\underline{L})| \leq H_{L, K, u}$$

for  $t \leq u$ . Taking expectations above and applying (4.13) and (4.15), one gets (4.14).  $\square$

Lemma 4.1 also gives an upper bound for the voter model on the probability of offspring of a parent at the origin escaping from  $B(L)$  by time  $u$ . The bound will be used in the proof of Lemma 4.3.

**COROLLARY 2.** *Suppose that  $L$  is large and  $L \leq u^{3/4}$ . Then there is a constant  $C$  so that*

$$P(\eta_t^{(0)} \not\subset B(L) \text{ for some } t \leq u) \leq CLu \exp(-L^2/8u).$$

**PROOF.** We use the duality relationship (2.11). From this, it follows that

$$\{\eta_t^{(0)} \not\subset B(L) \text{ for some } t \leq u\} \subset \{H_{L,K,u} > 0\}$$

for  $K = L/2\sqrt{u}$ . Consequently,

$$P(\eta_t^{(0)} \not\subset B(L) \text{ for some } t \leq u) \leq EH_{L,K,u}.$$

The statement then follows from (4.13).  $\square$

In the proofs of Lemma 4.2 and Propositions 2 and 3, we will use the following analog of (3.16). Here, we set

$$(4.16) \quad \bar{L} = L + 2\sqrt{u}(\log L)^{3/4}, \text{ where } u = L^2/(\log L)^2.$$

(The reason for this choice of  $\bar{L}$  will become clear in the proof of Lemma 4.2.) Using Lemma 3.2 again, this time with  $K = (\log L)^{3/4}$ , we see that for  $t \leq u$ ,

$$(4.17) \quad E|\zeta_t^{B(L)} \cap B(\bar{L})^c| \leq 8L^2 \exp(-(\log L)^{3/2}/2),$$

which tends to 0 as  $L \rightarrow \infty$ .

Lemma 4.2 gives asymptotics on  $E|\zeta_t^{B(L)}|$  for large  $t \leq L^2/(\log L)^2$ . It will be used in the proofs of both of the propositions.

**LEMMA 4.2.** *Let  $\gamma(L)$  be any function with  $\gamma(L) \rightarrow \infty$  as  $L \rightarrow \infty$ . For all  $\varepsilon > 0$ , if  $L$  is large enough, then*

$$(4.18) \quad 1 - \varepsilon \leq \frac{\pi t E|\zeta_t^{B(L)}|}{L^2 \log t} \leq 1 + \varepsilon \text{ for } t \in [\gamma(L), L^2/(\log L)^2].$$

**PROOF.** For the upper bound, we write

$$(4.19) \quad E|\zeta_t^{B(L)}| = E|\zeta_t^{B(L)} \cap B(\bar{L})| + E|\zeta_t^{B(L)} \cap B(\bar{L})^c|,$$

where  $\bar{L}$  is given by (4.16), and argue as in (3.13) and (3.14). Since  $\zeta_t^{B(L)} \subset \bar{\zeta}_t$ , one has

$$(4.20) \quad E|\zeta_t^{B(L)} \cap B(\bar{L})| \leq E|\bar{\zeta}_t \cap B(\bar{L})| \leq (\bar{L} + 1)^2 P(O \in \bar{\zeta}_t).$$

Our choice for  $\bar{L}$  implies that  $\bar{L}/L \rightarrow 1$  as  $L \rightarrow \infty$ . Together with (3.11), which gives the asymptotic behavior of  $P(O \in \bar{\zeta}_t)$ , this implies that the last term in (4.20) behaves asymptotically like  $L^2(\log t)/\pi t$  for large  $L$  and  $t$ . On the other hand, an upper bound for the last term in (4.19) is given by (4.17).



Together with the bound for (4.20) just derived, this gives the upper bound in (4.18).

For the lower bound in (4.18), we estimate  $E|\zeta_t^{B(L)} \cap B(\underline{L})|$ , where

$$(4.21) \quad \underline{L} = L - 2\sqrt{u}(\log L)^{3/4} \quad \text{and} \quad u = L^2/(\log L)^2.$$

By Corollary 1, for  $t \leq u$ ,

$$(4.22) \quad E|\bar{\zeta}_t \cap B(\underline{L})| - E|\zeta_t^{B(L)} \cap B(\underline{L})| \leq CLu \exp(-(\log L)^{3/2}/2).$$

This tends to 0 as  $L \rightarrow \infty$ . So, it suffices to analyze  $E|\bar{\zeta}_t \cap B(\underline{L})|$ . Using reasoning analogous to that just below (4.20), one can show that

$$(4.23) \quad E|\bar{\zeta}_t \cap B(\underline{L})| \geq (\underline{L} - 1)^2 P(O \in \bar{\zeta}_t) \sim L^2(\log t)/\pi t$$

for large  $L$  and  $t$ . Together with (4.22), this gives the lower bound in (4.18). □

Propositions 2 and 3 both involve extensions of Lemma 4.2. Proposition 2 asserts convergence in probability simultaneously over

$$t \in [\log L, L^2/(\log L)^2],$$

rather than just convergence of expectations. Proposition 3 extends the upper end of the range of  $t$  in Lemma 4.2, for the upper bound, to  $L^2/(\log \log L)^2$  from  $L^2/(\log L)^2$ . (The lower end of the range given here,  $\log L$ , is sufficient for our purposes.) Since Proposition 3 involves less work, we show it first.

PROOF OF PROPOSITION 3. On account of Lemma 4.2, it is enough to consider  $t \in [L^2/(\log L)^2, L^2/(\log \log L)^2]$ . We again use the decomposition

$$(4.24) \quad E|\zeta_t^{B(L)}| = E|\zeta_t^{B(L)} \cap B(M)| + E|\zeta_t^{B(L)} \cap B(M)^c|,$$

where we set

$$(4.25) \quad M = \bar{L} + 2\sqrt{v}(\log \log L)^{3/4} \quad \text{with} \quad v = L^2/(\log \log L)^2$$

and  $\bar{L}$  is given by (4.16). Our choice for  $M$  implies that  $M/L \rightarrow 1$  as  $L \rightarrow \infty$ . Reasoning as in Lemma 4.2 for the first term on the right-hand side of (4.24) gives

$$(4.26) \quad E|\zeta_t^{B(L)} \cap B(M)| \leq (M + 1)^2 P(O \in \bar{\zeta}_t) \sim L^2(\log t)/\pi t$$

for large  $L$  and  $t$ .

To demonstrate Proposition 3, we need to show that the second term on the right-hand side of (4.24) goes to 0. For this, we first consider  $\zeta_t^{B(L)}$  at the intermediate time  $u = L^2/(\log L)^2$ . By (4.17),

$$(4.27) \quad E|\zeta_u^{B(L)} \cap B(\bar{L})^c| \rightarrow 0 \quad \text{as} \quad L \rightarrow \infty.$$

In investigating  $E|\zeta_t^{B(L)} \cap B(M)^c|$  over  $[u, v]$ , we therefore only need to consider those particles inside  $B(\bar{L})$  at time  $u$ . Using Lemma 4.2,

$$(4.28) \quad E|\zeta_u^{B(L)} \cap B(\bar{L})| \leq E|\zeta_u^{B(L)}| \leq L^2(\log u)/u \leq 2(\log L)^3$$

for large  $L$ . These particles need to move a distance of at least  $\sqrt{v}(\log \log L)^{3/4}$  over  $t \in [u, v]$  to enter  $B(M)$ . By Lemma 3.2, with  $K = (\log \log L)^{3/4}$ , the probability that a given one of these particles does this by time  $v$  is at most  $8 \exp(-(\log \log L)^{3/2}/2)$ . With (4.28), this shows that the expected number of particles starting in  $B(\bar{L})$ , which visit  $B(M)$  at some time in  $[u, v]$ , is at most

$$16(\log L)^3 \exp(-(\log \log L)^{3/2}/2),$$

which approaches 0 as  $L \rightarrow \infty$ . Together with (4.27), the limit shows that

$$E|\zeta_t^{B(L)} \cap B(M)^c| \rightarrow 0$$

uniformly over  $t \in [u, v]$ . This demonstrates Proposition 3.  $\square$

To demonstrate Proposition 2, we need the following variance estimate. Note that the bound is smaller than  $(E|\zeta_t^{B(L)}|)^2$  by a factor of  $(\log L)^{-3/2}$ .

LEMMA 4.3. *There is a constant  $C$  so that if  $L$  is large, then for all  $t \in [\log L, L^2/(\log L)^3]$ ,*

$$(4.29) \quad \text{var}(|\zeta_t^{B(L)} \cap B(2L)|) \leq C(\log L)^{-3/2} \left( \frac{L^2 \log t}{t} \right)^2.$$

PROOF. Let  $F_x$  denote the event that  $x \in \zeta_t^{B(L)}$ . Writing  $|\zeta_t^{B(L)} \cap B(2L)|$  as a sum of the indicators of  $F_x$ , it is easy to see that

$$(4.30) \quad \text{var}(|\zeta_t^{B(L)} \cap B(2L)|) = \sum_{x, y \in B(2L)} (P(F_x \cap F_y) - P(F_x)P(F_y)).$$

Let  $\eta_t^{(x)}$  denote the voter model started from  $\{x\}$ . We recall from Section 2 that, by duality,  $F_x$  is also the event that  $\eta_t^{(x)} \cap B(L) \neq \emptyset$ . Using the same percolation substructure,  $F_x \cap F_y$  is the event that if we start the multitype voter model with, say, opinion 1 at  $x$ , opinion 2 at  $y$  and opinion 0 elsewhere, then opinions 1 and 2 are both present in  $B(L)$  at time  $t$ . Let  $G(x, y)$  denote the event that  $\eta_s^{(x)}$  and  $\eta_s^{(y)}$  intersect at some time  $s \leq t$ . By using two independent graphical representations to construct  $\eta_s^{(x)}$  and  $\eta_s^{(y)}$  until the first time they intersect, it is not difficult to check that the summands on the right-hand side of (4.30) are bounded above by  $P(G(x, y))$ . For more details see the proof of (2.6) in Griffeath (1979).

We now proceed to estimate  $G(x, y)$ . For this, we first note that taking  $u = t$  and  $L = 2(\log t)^{2/3}\sqrt{t}$  in Corollary 2 gives that

$$(4.31) \quad P(\eta_s^{(0)} \not\subset B((2 \log t)^{2/3} \sqrt{t}) \text{ for some } s \leq t) \leq C(\log t)^{2/3} t^{3/2} \exp(-(\log t)^{4/3}/2) \leq t^{-4}$$

for large enough  $t$ . We define  $l$  so that

$$(4.32) \quad t = l^2/(\log l)^{3/2}.$$

(Assume that  $l \geq e^{3/4}$  so that the choice is unique.) Since  $\log t \leq 2 \log l$ , it is easy to see that

$$(\log t)^{2/3} \sqrt{t} \leq (2 \log l)^{2/3} \frac{l}{(\log l)^{3/4}} = o(l).$$

In particular,

$$(4.33) \quad l > 2(\log t)^{2/3} \sqrt{t}$$

for  $l$  chosen large enough. It follows from (4.31) and (4.33) that for  $x, y \in \mathbf{Z}^2$  with  $\|x - y\| > l$  and  $l$  large,

$$(4.34) \quad \begin{aligned} P(G(x, y)) &\leq P(\zeta_s^{(x)} \not\subset x + B(2(\log t)^{2/3} \sqrt{t}) \text{ or} \\ &\zeta_s^{(y)} \not\subset y + B(2(\log t)^{2/3} \sqrt{t}) \text{ for some } s \leq t) \\ &\leq 2t^{-4}. \end{aligned}$$

We analyze the terms on the right-hand side of (4.30) by dividing them up according to whether (a)  $\|x - y\| > l$  or (b)  $\|x - y\| \leq l$ . In the first case, we employ (4.34). Since there are at most  $(2L + 1)^4$  pairs  $(x, y)$  with  $x, y \in B(2L)$ ,

$$(4.35) \quad \begin{aligned} &\sum_{\substack{x, y \in B(2L) \\ \|x - y\| > l}} (P(F_x \cap F_y) - P(F_x)P(F_y)) \\ &\leq \sum_{\substack{x, y \in B(2L) \\ \|x - y\| > l}} P(G(x, y)) \leq 2t^{-4}(2L + 1)^4 \\ &\leq C(\log L)^{-2} \frac{L^4(\log t)^2}{t^2} \end{aligned}$$

for large  $L$  and  $t \geq \log L$ .

To demonstrate (4.29), it therefore suffices to show that

$$(4.36) \quad \sum_{\substack{x, y \in B(2L) \\ \|x - y\| \leq l}} P(F_x \cap F_y) \leq C(\log L)^{-3/2} \left( \frac{L^2 \log t}{t} \right)^2.$$

We use the bound

$$(4.37) \quad P(F_x \cap F_y) \leq P(\eta_t^{(x)} \neq \emptyset, \eta_t^{(y)} \neq \emptyset).$$

By Lemma 1 of Arratia (1981), for all  $x, y$ , and  $t$ ,

$$(4.38) \quad P(\eta_t^{(x)} \neq \emptyset, \eta_t^{(y)} \neq \emptyset) \leq (P(\eta_t^{(0)} \neq \emptyset))^2.$$

By (2.2) and (3.11),

$$P(\eta_t^{(0)} \neq \emptyset) = P(O \in \bar{\xi}_t) \sim \frac{\log t}{\pi t}$$

for large  $t$ . So, by (4.37) and (4.38),

$$(4.39) \quad P(F_x \cap F_y) \leq \left(\frac{\log t}{t}\right)^2$$

for large  $t$  and all  $x$  and  $y$ . There are at most  $(2l + 1)^2(2L + 1)^2 \leq Cl^2L^2$  pairs of sites  $(x, y)$  with  $x, y \in B(2L)$  and  $\|x - y\| \leq l$ . It therefore follows from (4.39) that for large  $L$  and  $t \geq \log L$ ,

$$(4.40) \quad \sum_{\substack{x, y \in B(2L) \\ \|x - y\| \leq l}} P(F_x \cap F_y) \leq C \left(\frac{\log t}{t}\right)^2 l^2 L^2 \leq C(\log t)^{5/2} \frac{L^2 \log t}{t}.$$

To convert the bound into the form given in (4.36), we note that  $(\log t)/t$  is decreasing over  $t \in [e, L^2/(\log L)^3]$ . For such  $t$  and  $L \geq (\log L)^3$ , one has

$$(\log L)^4 \leq \frac{2L^2 \log t}{t}.$$

It follows that

$$(4.41) \quad (\log t)^{5/2} \leq 2^{5/2}(\log L)^{5/2} \leq 2^{5/2}(\log L)^{-3/2} \frac{2L^2 \log t}{t}.$$

Plugging (4.41) into (4.40) gives (4.36) for large  $L$  and  $t \in [\log L, L^2/(\log L)^3]$ . This in turn implies (4.29).  $\square$

The estimates for  $E|\zeta_t^{B(L)}|$  in Lemma 4.2 and  $\text{var}(|\zeta_t^{B(L)} \cap B(2L)|)$  in Lemma 4.3, together with Chebyshev's inequality, lead easily to a proof of Proposition 2.

PROOF OF PROPOSITION 2. Set  $\check{\zeta}_t^{B(L)} = \zeta_t^{B(L)} \cap B(2L)$  and fix  $\delta \in (0, 1)$ . Application of Chebyshev's inequality to (4.29) implies that there is a constant  $C$  so that

$$(4.42) \quad P\left(\left||\check{\zeta}_t^{B(L)}| - E|\check{\zeta}_t^{B(L)}|\right| > \frac{\delta L^2 \log t}{\pi t}\right) \leq \frac{C}{\delta^2(\log L)^{3/2}}$$

for large  $L$  and  $t \in [\log L, L^2/(\log L)^3]$ . Since  $\bar{L} \leq 2L$ , (4.17) implies that

$$(4.43) \quad E|\zeta_t^{B(L)}| - E|\check{\zeta}_t^{B(L)}| \leq 8L^2 \exp(-(\log L)^{3/2}/2),$$

which approaches 0 as  $L \rightarrow \infty$ . Moreover, since  $|\check{\zeta}_t^{B(L)}| \leq |\zeta_t^{B(L)}|$ , one can apply Chebyshev's inequality to (4.43) to bound  $|\zeta_t^{B(L)}| - |\check{\zeta}_t^{B(L)}|$  as well. It therefore follows from (4.42) and (4.43) that, for a slightly larger  $C$ ,

$$(4.44) \quad P\left(\left||\zeta_t^{B(L)}| - E|\zeta_t^{B(L)}|\right| > \frac{2\delta L^2 \log t}{\pi t}\right) \leq \frac{C}{\delta^2(\log L)^{3/2}}$$

for  $t \in [\log L, L^2/(\log L)^3]$ .

Since  $|\zeta_t^{B(L)}|$  is decreasing in  $t$ , it suffices to control its behavior at the points  $t_j$  given by

$$(4.45) \quad t_j = \frac{\log L}{(1 - \delta)^j} \wedge \frac{L^2}{(\log L)^3}, \quad j = 0, 1, \dots, J,$$

where  $J$  is the first value of  $j$  at which  $t_j = L^2/(\log L)^3$ . From

$$\frac{\log L}{(1 - \delta)^{J-1}} < \frac{L^2}{(\log L)^3},$$

one obtains

$$(4.46) \quad J < \frac{-2 \log L}{\log(1 - \delta)} + 1 < \frac{2 \log L}{\delta} + 1.$$

Substitution of  $t_0, t_1, \dots, t_J$  into (4.44), together with (4.46), gives

$$(4.47) \quad P\left(\left| |\zeta_{t_j}^{B(L)}| - E|\zeta_{t_j}^{B(L)}| \right| \leq \frac{2\delta L^2 \log t_j}{\pi t_j} \text{ for all } j \leq J\right) \rightarrow 1$$

as  $L \rightarrow \infty$ . Lemma 4.2 gives asymptotics for the mean. Together with the monotonicity of  $|\zeta_t^{B(L)}|$ , it implies that

$$\begin{aligned} P\left( (1 - 3\delta) \frac{L^2 \log t_{j+1}}{\pi t_{j+1}} \leq |\zeta_t^{B(L)}| \right. \\ \left. \leq (1 + 3\delta) \frac{L^2 \log t_j}{\pi t_j} \text{ for } t \in [t_j, t_{j+1}], j \leq J \right) \rightarrow 1 \end{aligned}$$

as  $L \rightarrow \infty$ . Using the definition of  $t_j$ , one obtains that

$$P\left( (1 - 3\delta)(1 - \delta) \leq \frac{\pi t |\zeta_t^{B(L)}|}{L^2 \log t} \leq \frac{1 + 3\delta}{1 - \delta} \text{ for all } t \in [\log L, L^2/(\log L)^3] \right) \rightarrow 1$$

as  $L \rightarrow \infty$ . The bound (4.1) follows by setting  $\delta = \varepsilon/5 \wedge 1/5$ .  $\square$

**5. Proof of Theorem 3.**

PROOF OF THE UPPER BOUND. We compute the upper bound in (1.3) for  $0 \leq r < 1$ . As in Section 4, we employ the duality relationships in Section 2 for the process  $\xi_t$  started in equilibrium. Let  $0 < \varepsilon < (1 - r) \wedge \frac{1}{3}$ , and  $T_0 = 0$ ,  $T_1 = L^{2r(1-\varepsilon)}$ ,  $T_2 = L^{2(1-\varepsilon)}$  and  $T_3 = L^2$ . Let  $y_i$  be the expected number of mutations along the extended ancestral path  $\zeta_s^{B(L)}$  between times  $T_{i-1}$  and  $T_i$ . That is,  $y_i = EY_{T_{i-1}, T_i}^{B(L)}$ . Using (2.7), one has

$$(5.1) \quad EN_r \leq y_1 + y_2 + z_{T_2}^{B(L)}.$$

We proceed to estimate the terms on the right-hand side.

For  $y_1$ , note that by Lemma 3.1,

$$(5.2) \quad z_t^{B(L^r)} \leq CL^{2r}(\log t)/t$$

for  $e^2 \leq t \leq L^{2r(1-\varepsilon)}$ . Using (2.8) and integrating as in (3.7), we get that

$$y_1 \leq C\alpha L^{2r} \left( e^2 + \int_{e^2}^{T_1} \frac{\log t}{t} dt \right) \leq CL^{2r-2}(\log L)^2,$$

where, as before, we change  $C$  when convenient. Since  $r < 1$ , it follows that

$$(5.3) \quad y_1 \rightarrow 0 \quad \text{as } L \rightarrow \infty.$$

To estimate  $y_2$ , note that for  $t \geq T_1$ ,

$$z_t^{B(L^r)} \leq z_{T_1}^{B(L^r)}.$$

So, by using (2.8) and (5.2), it follows that

$$(5.4) \quad y_2 \leq \alpha(T_2 - T_1)z_{T_1}^{B(L^r)} \leq CL^{-2\varepsilon(1-r)} \log L \rightarrow 0$$

as  $L \rightarrow \infty$ .

To estimate  $z_{T_2}^{B(L^r)}$ , we employ the proposition on page 615 of Bramson, Cox and Griffeath (1986). First note that there is nothing to prove when  $r = 0$ , so we can suppose  $r > 0$ . Setting  $t = L^r/2$  and  $\beta = 2(1 - \varepsilon)/r$ , we have

$$(5.5) \quad z_{T_2}^{B(L^r)} \leq A \left[ \left( \frac{2(1 - \varepsilon)}{r} - 2 \right)^{-1} \vee 1 \right]$$

for large  $L$ , where  $A$  is a fixed constant not dependent on  $\varepsilon$ . Together, (5.1), (5.3), (5.4) and (5.5) imply that

$$(5.6) \quad \limsup_{L \rightarrow \infty} EN_r \leq A \left[ \left( \frac{2(1 - \varepsilon)}{r} - 2 \right)^{-1} \vee 1 \right].$$

Letting  $\varepsilon \rightarrow 0$ , one obtains from (5.6) that

$$(5.7) \quad \limsup_{L \rightarrow \infty} EN_r \leq A \left[ \left( \frac{2}{r} - 2 \right)^{-1} \vee 1 \right] = \begin{cases} 1, & \text{for } 0 \leq r \leq \frac{2}{3}, \\ \frac{Ar}{2(1-r)}, & \text{for } \frac{2}{3} \leq r < 1. \end{cases}$$

This gives the upper bound in (1.3).  $\square$

**PROOF OF CONVERGENCE IN DISTRIBUTION.** By (5.3) and (5.4),  $y_1 + y_2 \rightarrow 0$  as  $L \rightarrow \infty$ . This says that for large  $L$  and fixed  $\varepsilon > 0$ , there are typically no mutations before time  $T_2$  along the extended ancestral path. So, in the limit,  $|\zeta_{T_2}^{B(L^r)}|$  provides an upper bound for  $N_r$ . Turning again to Bramson, Cox and Griffeath (1986), we have, for  $0 < r < r'$ ,

$$(5.8) \quad \lim_{L \rightarrow \infty} P(|\zeta_{L^{2r'}}^{B(L^r)}| = k) = p_{\infty, k}(r/r').$$

This equation is a paraphrase of Theorem 1 of Bramson, Cox and Griffeath (1986), where one sets  $t = L^{2r'}$  and  $\alpha = r/r'$ . (This  $\alpha$  is not the mutation rate.) The terms  $p_{\infty, k}(\alpha)$  are equal to  $P(D_{\log(1/\alpha)} = k)$ , where  $D_t$  is the pure death chain starting from  $\infty$  that jumps from  $j$  to  $j - 1$  at rate  $\binom{j}{2}$ . [See (3.10) on page 357 of Cox and Griffeath (1986) for details.] Consequently, for any  $k$ ,

$$(5.9) \quad \liminf_{L \rightarrow \infty} P(N_r \leq k) \geq \lim_{L \rightarrow \infty} P(|\zeta_{T_2}^{B(L^r)}| \leq k) = \sum_{j=0}^k p_{\infty, j}(r/(1 - \varepsilon)).$$

For the corresponding inequality in the opposite direction, we note that by (5.5), the expected number of particles on the extended ancestral path at time  $T_2$  is bounded. Therefore, with probability close to 1 for large  $L$ , all extant particles will experience mutations by time  $T_3$ . In particular, by (2.3) and (2.6),  $|\zeta_{T_3}^{B(L^r)}|$  provides a lower bound for  $N_r$  in the limit. Thus, using (5.8),

$$(5.10) \quad \limsup_{L \rightarrow \infty} P(N_r \leq k) \leq \lim_{L \rightarrow \infty} P(|\zeta_{T_3}^{B(L^r)}| \leq k) = \sum_{j=0}^k p_{\infty, j}(r).$$

Since the functions  $p_{\infty, j}(\cdot)$  are continuous, (5.9) and (5.10) imply

$$(5.11) \quad \lim_{L \rightarrow \infty} P(N_r = k) = p_{\infty, k}(r).$$

This identifies the limiting distribution in Theorem 3.  $\square$

PROOF OF THE LOWER BOUND. We still need to demonstrate the lower bound in (1.3). Fatou's lemma and (5.11) imply

$$\liminf_{L \rightarrow \infty} EN_r \geq \sum_{k=1}^{\infty} k p_{\infty, k}(r) = ED_{\log(1/r)}.$$

Since  $\log(1/r) \sim 1 - r$  as  $r \rightarrow 1$ , it suffices to show the following lemma.

LEMMA 5.1. As  $\varepsilon \rightarrow 0$ ,  $\varepsilon ED_\varepsilon \rightarrow 2$ .

PROOF. The time  $t_n$  for  $D_t$  to make its transition from  $n + 1$  to  $n$  is exponential with mean  $2/n(n + 1)$ . Let  $\tau_n = t_n + t_{n+1} + \dots$  be the time to reach  $n$  starting from  $\infty$ . It is easy to see that

$$E\tau_n = \sum_{m=n}^{\infty} \frac{2}{m(m + 1)} = \sum_{m=n}^{\infty} \left( \frac{2}{m} - \frac{2}{m + 1} \right) = \frac{2}{n},$$

$$\text{var}(\tau_n) = \sum_{m=n}^{\infty} \frac{4}{m^2(m + 1)^2} \leq \int_{n-1}^{\infty} \frac{4}{x^4} dx = \frac{4}{3(n - 1)^3}.$$

Thus,  $n\tau_n$  has mean 2 and variance at most  $4n^2/3(n - 1)^3 \leq 8/n$  for  $n \geq 3$ , and Chebyshev's inequality implies that  $n\tau_n \rightarrow 2$  in probability as  $n \rightarrow \infty$ . Since

$$P(\varepsilon D_\varepsilon \leq x0) = P(\tau_{\lceil x/\varepsilon \rceil} \leq \varepsilon),$$

it follows easily that  $\varepsilon D_\varepsilon \rightarrow 2$  in probability as  $\varepsilon \rightarrow 0$ .

To get  $\varepsilon ED_\varepsilon \rightarrow 2$ , it suffices to prove uniform integrability, for example, that  $\varepsilon^2 ED_\varepsilon^2$  or, equivalently,  $\int_0^\infty xP(\varepsilon D_\varepsilon > x) dx$  remains bounded as  $\varepsilon \rightarrow 0$ . For  $\varepsilon < 1$  and  $x > 4$ , we have

$$\begin{aligned} P(\varepsilon D_\varepsilon > x) &= P\left(\left[\frac{x}{\varepsilon}\right] \tau_{[x/\varepsilon]} > \varepsilon \left[\frac{x}{\varepsilon}\right]\right) \leq \frac{\text{var}([x/\varepsilon] \tau_{[x/\varepsilon]})}{(\varepsilon[x/\varepsilon] - 2)^2} \\ &\leq \frac{16\varepsilon}{x} \frac{1}{(x/2 - 2)^2} = \frac{64\varepsilon}{x(x-4)^2}. \end{aligned}$$

Thus,

$$\int_0^\infty xP(\varepsilon D_\varepsilon > x) dx \leq 25 + \int_5^\infty \frac{64\varepsilon}{(x-4)^2} dx \leq 25 + 64\varepsilon$$

and we are done.  $\square$

## REFERENCES

- ARRATIA, R. (1981). Limiting point processes for rescalings of coalescing and annihilating random walks on  $\mathbf{Z}^d$ . *Ann. Probab.* **9** 909–936.
- ARRHENIUS, O. (1921). Species and area. *Journal of Ecology* **9** 95–99.
- 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.
- BRAMSON, M. and GRIFFEATH, D. (1979). Renormalizing the three-dimensional voter model. *Ann. Probab.* **7** 418–432.
- BRAMSON, M. and GRIFFEATH, D. (1980). Asymptotics for some interacting particle systems on  $\mathbf{Z}^d$ . *Z. Warsch. Verw. Gebiete* **53** 183–196.
- CLIFFORD, P. and SUDBURY, A. (1973). A model for spatial conflict. *Biometrika* **60** 681–588.
- COLEMAN, B. D. (1982). On random placement and species-area-relations. *Math. Biosci.* **54** 191–215.
- CONNOR, E. F. and MCCOY, E. D. (1979). The statistics and biology of the species-area relationship. *American Naturalist* **113** 791–833.
- COX, J. T. and GRIFFEATH, D. (1986). Diffusive clustering in the two dimensional voter model. *Ann. Probab.* **14** 347–370.
- DONNELLY, P. and TAVARÉ, S. (1995). Coalescents and genealogical structure under neutrality. *Annual Review of Genetics* **29** 401–421.
- DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth, Belmont, CA.
- DURRETT, R. and LEVIN, S. (1996). Spatial models for species area curves. *J. Theoret. Biol.* To appear.
- EWENS, W. (1990). Population genetics—the past and the future. In *Mathematical and Statistical Developments of Evolutionary Theory* (S. Lessard, ed.) 177–277. Kluwer, Dordrecht.
- FISHER, R. A., CORBET, A. S. and WILLIAMS, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* **12** 42–58.
- GRIFFEATH, D. (1979). *Additive and Cancellative Interacting Particle Systems. Lecture Notes in Math.* **724**. Springer, Berlin.
- HOLLEY, R. A. and LIGGETT, T. M. (1975). Ergodic theorems for weakly interacting systems and the voter model. *Ann. Probab.* **3** 643–663.
- HUBBELL, S. P. (1992). Speciation, dispersal, and extinction: an equilibrium theory of species-area relationships. Preprint.



- KARLIN, S. (1967). Central limit theorems for certain infinite urn schemes. *Journal of Mathematical Mechanics* **17** 373–401.
- KREITMAN, M. and AKASHI, H. (1995). Molecular evidence for natural selection. *Annual Review of Ecological Systems* **26** 403–422.
- LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.
- MACARTHUR, R. H. and WILSON, E. O. (1967). *The Theory of Island Biogeography*. Princeton Univ. Press.
- NEUHAUSER, C. (1992). Ergodic theorems for the multitype contact process. *Probab. Theory Related Fields* **91** 467–506.
- PITMAN, J. (1996). Species sampling models. Unpublished manuscript.
- PRESTON, F. W. (1962). The canonical distribution of commonness and rarity. I, II. *Ecology* **43** 185–215; 410–432.
- ROUAULT, A. (1978). Lois de Zipf et sources markoviennes. *Ann. Inst. H. Poincaré Probab. Statist.* **14** 169–178.
- WATSON, H. C. (1835). *Remarks on the Geographical Distribution of British Plants*. Longmans, London.
- WILLIAMSON, M. (1988). Relationship of species number to area, distance and other variables. In *Analytical Biogeography* (A. A. Myers and P. S. Giller, eds.) Chap. 4. Chapman and Hall, London.

MAURY BRAMSON  
DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF WISCONSIN  
MADISON, WISCONSIN 53706  
E-MAIL: bramson@math.wisc.edu

J. THEODORE COX  
DEPARTMENT OF MATHEMATICS  
SYRACUSE UNIVERSITY  
SYRACUSE, NEW YORK 13244  
E-MAIL: jtcox@gumby.syr.edu

RICHARD DURRETT  
DEPARTMENT OF MATHEMATICS  
WHITE HALL  
CORNELL UNIVERSITY  
ITHACA, NEW YORK 14853  
E-MAIL: rtd1@cornell.edu