The Annals of Probability
1986, Vol. 14, No. 2, 347–370

SPECIAL INVITED PAPER

DIFFUSIVE CLUSTERING IN THE TWO DIMENSIONAL VOTER MODEL

BY J. THEODORE COX[1] AND DAVID GRIFFEATH[2]

Syracuse University and University of Wisconsin

We study the behavior of an interacting particle system known as the voter model in two dimensions. This process provides a simple example of "critical clustering" among two colors, say green and black, in the plane. The paper begins with some computer simulations, and a survey of known results concerning the voter model in the three qualitatively distinct cases: three or more dimensions (high), one dimension (low), and two dimensions (critical). Our main theorem, for the planar model, states roughly that at large times t the proportion of green sites on a box of side $t^{\alpha/2}$ centered at the origin fluctuates with α according to a time change of the Fisher–Wright diffusion. Some applications of the theorem, and several related results, are described.

1. Introduction. Start by coloring each site of the two dimensional integer lattice $\mathbb{Z}^2$ either green or black at random. Now consider a stochastic process which evolves according to the following very simple dynamic: At any time $t \geq 0$ each site waits an exponential holding time with mean 1, chooses a neighboring site (distance 1 away) at random, and gives that site its color. We have written a computer simulation of this evolution on a 512×256 box with periodic boundary conditions and about 5,000 changes of state per second. Figures 1–3 show the "movie" at various times.

The pictures clearly show that our process, known as the two dimensional voter model, *clusters* as time goes on. The central objective of this paper is to formulate and prove some precise mathematical statements concerning the clustering. Perhaps the most basic question to answer is: How big are the clusters at a large time t? We will prove a theorem which effectively answers this question, and gives a good deal of quantitative information about the spatial correlations.

The organization of the paper is as follows. Section 2 contains a survey of known results concerning the voter model. The dynamics described above define a stochastic process on the integer lattice $\mathbb{Z}^d$ for any dimension d, and it turns out that the behavior is qualitatively different in the three cases $d = 1$, $d = 2$, and $d \geq 3$. We review these distinctions in order to illuminate the "critical"

Received April 1985; revised August 1985.
[1] Partially supported by the National Science Foundation under Grant DMS-841317.
[2] Partially supported by the National Science Foundation under Grant DMS-830549.
AMS 1980 *subject classification*. Primary 60K35.
Key words and phrases. Infinite particle system, voter model, coalescing random walks, genetics diffusion, clustering, exchangeable random field.

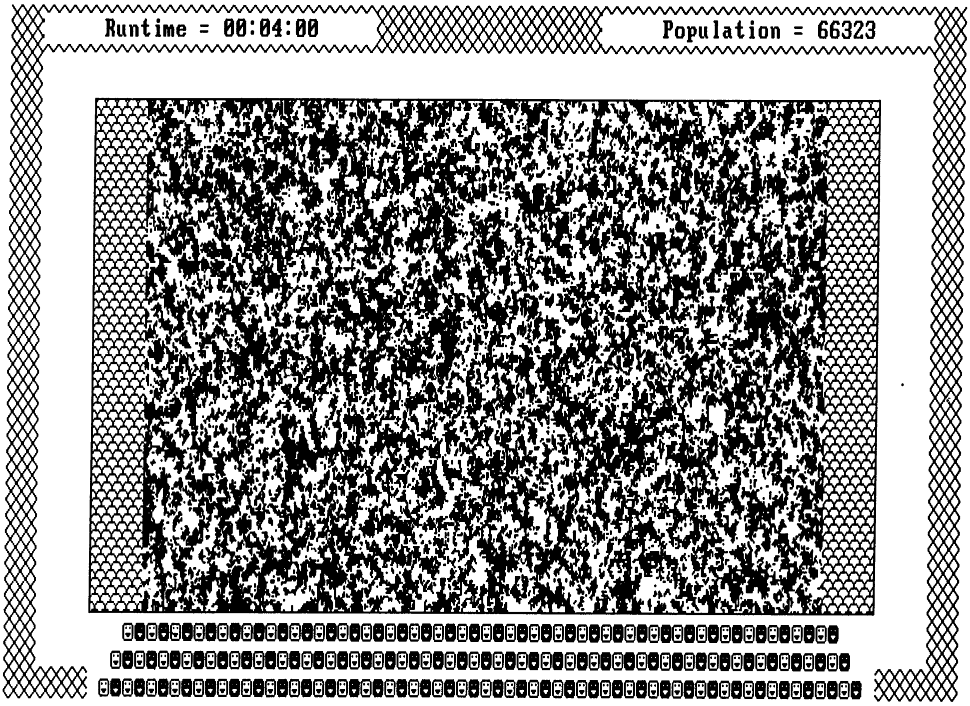


FIG. 1. *A planar voter model simulation after 4 min. The "population" is the number of black sites. The flip rate is roughly 30 sec / site. The movie, written in assembly language for the IBM-PC with hi-res monochrome graphics, is available from the authors.*

nature of the two dimensional model. Next, in Section 3 we present some new results, culminating with Theorem 5. For readability, only the major steps of the proofs are presented there. Section 4 describes some applications of the main theorem as well as various extensions and related results which can be proved using the same techniques; details will appear elsewhere. Finally, the technicalities of our proofs are relegated to Section 5.

2. Background. Several different interacting systems have been proposed as simple schemes to represent the random evolution of opinions (e.g., "for Dewey," "for Truman") in a spatial community, where each individual is influenced by neighbors. Mathematically expedient assumptions are that each opinion changes at a rate proportional to the number of neighbors who disagree, and that the dynamics are symmetric in the two opinions (i.e., unbiased). A finite lattice process of this sort recently percolated into the April 1985 issue of *Scientific American* (see page 26, complete with elephants and donkeys). Early references to similar models in the mathematical genetics literature are [20], [21], and [24].

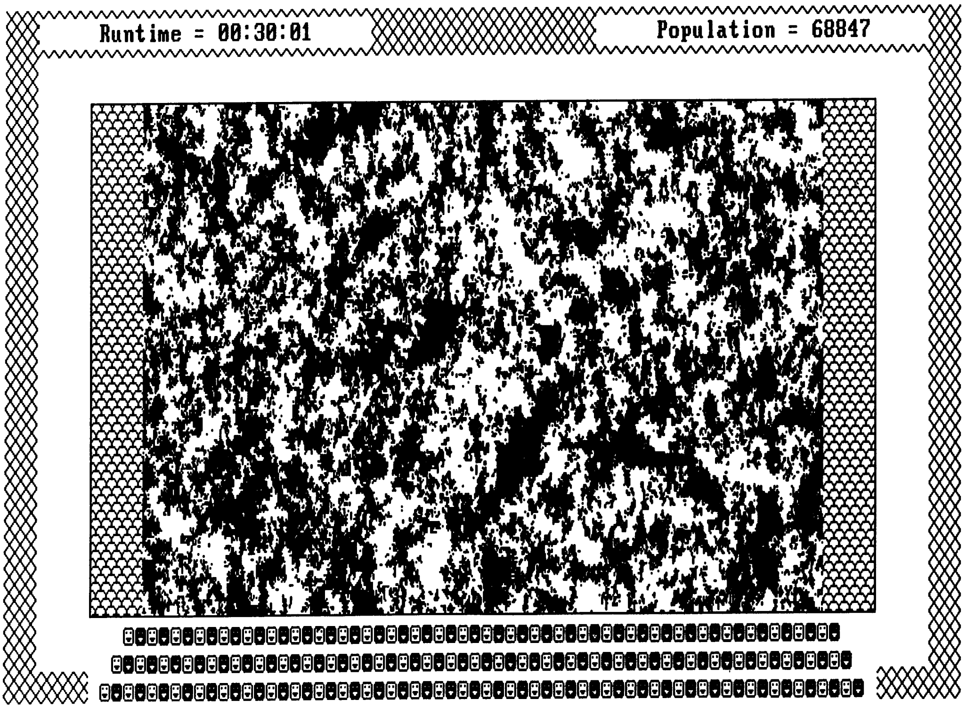


FIG. 2. *Same as Fig. 1 for 30 min.*

The first rigorous analysis of the precise system we will treat here was by Clifford and Sudbury [8]. They studied the infinite lattice version, calling it the invasion process to reflect the essential feature that regions of one opinion (or color) can only be penetrated by the other at the boundary. They also proved the most basic theorem about the voter model: that it is stable in three or more dimensions but clusters in one and two dimensions. To formulate this result precisely we need a little notation.

The voter model $\{\eta_t\}_{t \geq 0}$ is the Markov process on $\{0,1\}^{\mathbb{Z}^d}$ with rates specified by

$$\eta_t(x) \to 1 - \eta_t(x) \quad \text{at rate} \quad (2d)^{-1}\#\{y: |x-y| = 1, \eta_t(y) \neq \eta_t(x)\}.$$

Thus $\eta_t(x)$ codes the color (opinion, etc.) at site x at time t as a 0 or 1. Liggett [23] is an excellent source for a complete technical description of η_t, and for general background on infinite particle systems. To keep matters simple we will always assume that $\eta_t = \eta_t^\theta$ has initial distribution μ_θ, $0 < \theta < 1$, where μ_θ is product measure with density θ, i.e., $\mu_\theta\{\eta(x) = 1\} = \theta$ for all $x \in \mathbb{Z}^d$. The fundamental behavior of the voter model, discovered independently by Clifford and Sudbury [8] and Holley and Liggett [16], is described by

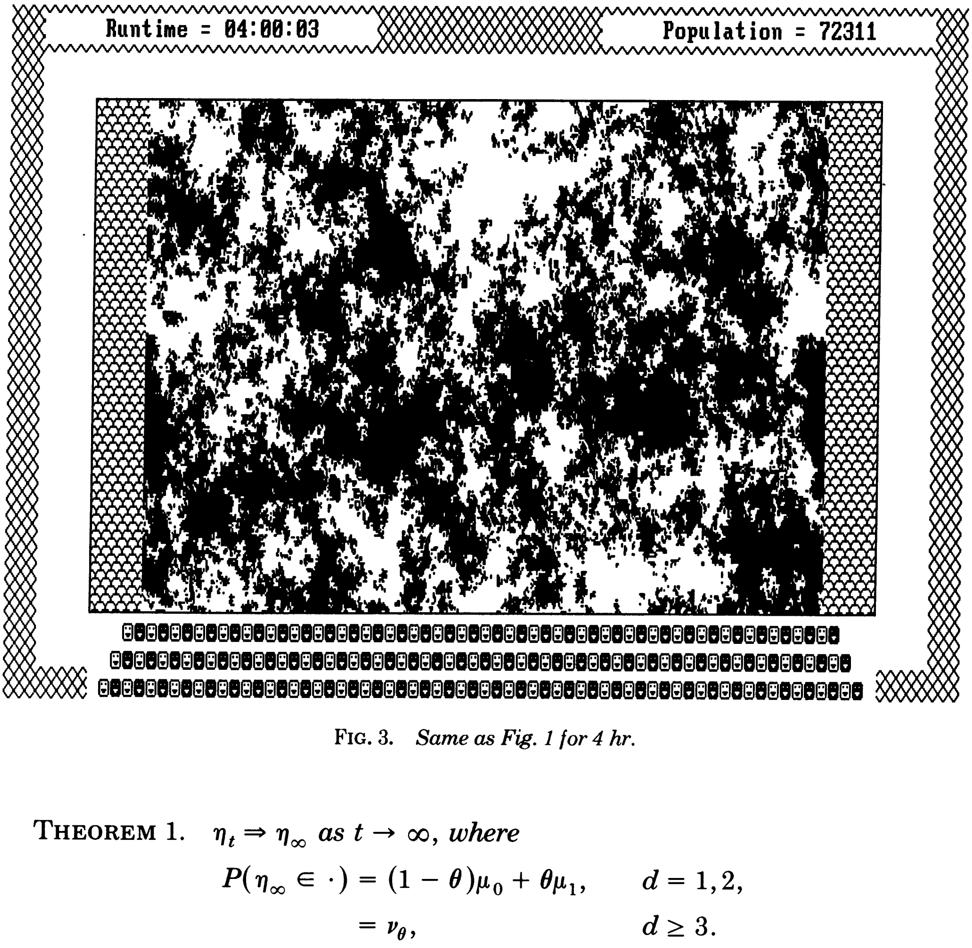


FIG. 3. *Same as Fig. 1 for 4 hr.*

THEOREM 1. $\eta_t \Rightarrow \eta_\infty$ *as* $t \to \infty$, *where*

$$P(\eta_\infty \in \cdot) = (1-\theta)\mu_0 + \theta\mu_1, \qquad d = 1,2,$$
$$= \nu_\theta, \qquad d \geq 3.$$

Here $\Rightarrow$ denotes weak convergence, μ_0 and μ_1 concentrate on all 0's and all 1's, respectively, and the ν_θ are neither independent nor totally correlated. Theorem 1 asserts that η_t approaches a nontrivial equilibrium if $d \geq 3$, whereas η_t approaches complete consensus for $d = 1,2$. Consensus means that for any $x, y \in \mathbb{Z}^d$,

$$\lim_{t\to\infty} P(\eta_t(x) \neq \eta_t(y)) = 0, \qquad d = 1,2, \tag{2.1}$$

indicating the formation of larger and larger *clusters* over time. Locally the configuration at large times is approximated by the flip of a $(\theta, 1-\theta)$ coin: Put down all 1's with probability θ, all 0's otherwise. Holley and Liggett coined the term "voter model," and fully determined the ergodic theory of the system in their beautiful paper [16]. Letting I_E denote the set of extreme invariant measures for η_t, they showed that $I_E = \{\nu_\theta,\ 0 \leq \theta \leq 1\}$ if $d \geq 3$, whereas $I_E = \{\mu_0, \mu_1\}$ if $d = 1$ or 2. They also described the domains of attraction of the equilibria. In addition, they considered a much more general class of voter models, one for each transition function $p(x, y)$. As the reader has probably guessed, transience or recurrence (of a symmetrization of p) determines whether a given model is stable or clusters. For simplicity we will treat only the uniform nearest neighbor case here; details of the more general framework may be found in [16] or [23] for example.

In three or more dimensions, when the voter model is stable, interest centers on the nature of the extreme invariant measures ν_θ, which exhibit unusual long range correlations. Let η_∞ denote the random field with measure ν_θ, and consider the centered block sum

$$S_n = \sum_{\|x\| \leq n} [\eta_\infty(x) - \theta] \qquad (\|x\| = \max\{x_i\}).$$

It turns out that to get a central limit theorem for S_n one needs to normalize by $n^{(d+2)/2}$ instead of the usual $n^{d/2}$. Moreover, if we define $S_n(x)$ in terms of the block sum of side $2n$ centered at $2nx$ (x a vector with integer coordinates), then this field of block sums, normalized by $n^{(d+2)/2}$, converges to a limiting Gaussian random field with covariance:

$$E[\eta_\infty(x)\eta_\infty(y)] = C_{d,\theta}\int_{B_x}\int_{B_y} \frac{1}{|u - v|^{d-2}}\, du\, dv, \tag{2.2}$$

where B_z is the cube of side 2 in $\mathbb{R}^d$ centered at z. See [4] for the proof. This result provides a particularly simple example of a strongly correlated equilibrium field for which one can compute the renorm limit of Wilson and Kadanoff. A good reference on renormalization is Sinai's book [27]. Holley and Stroock [17] have found a nice invariance principle corresponding to (2.2). They consider generalized random fields of renormalized sums at finite times, with centering and normalization by $t^{(d+2)/2}$ at time st^2. For $d \geq 3$, when $t \to \infty$ they get a limiting process in s which is generalized Ornstein–Uhlenbeck. As one would expect, the Gaussian field with covariances (2.2) is the equilibrium for their process.

Let us now turn to one dimension. The voter model on $\mathbb{Z}$ has been studied in considerable detail. The clustering which takes place is easy to conceptualize and to analyse mathematically, due to the linear nature of the lattice. In particular it is easy to see that the edges between clusters of opposite color execute simple random walks, with annihilation when two walks meet. (Annihilation occurs when a cluster has been "swallowed" by the surrounding pair of clusters so that these edges cease to exist.) A computer simulation illustrates this process nicely. With time running down the page (from $t = 0$ to $t = 347$) and using 720 lattice points, a space–time picture looks something like Figure 4.

One way to quantify the clustering we see there is to scale space appropriately with time. For the voter model on the integers this approach was initiated in [5], and fully realized by Arratia in his thesis [2]. Arratia's work shows that at large times t the one dimensional process consists of blocks of length $\sqrt{t}$. With appropriate normalization, the boundaries between these blocks converge as processes to annihilating Brownian motions (one motion starting at *every* real position). Limit theorems for the voter model can be read from this invariance

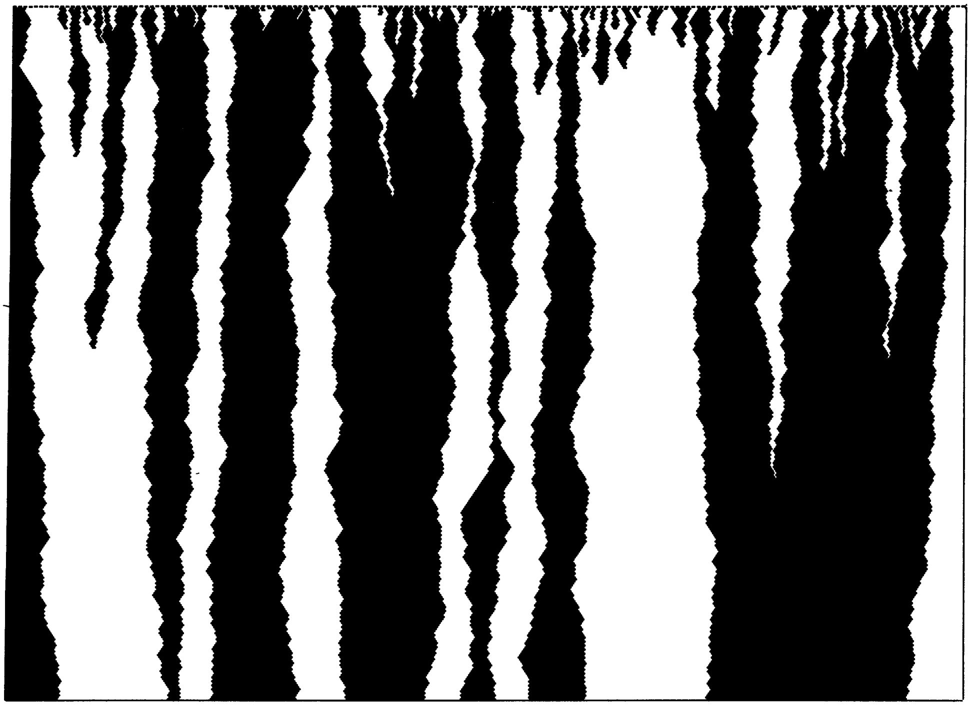


FIG. 4. *A simulation of the one dimensional voter model (space–time).*

principle. Here are some examples of the results which can be proved using the techniques of [2] and [5].

THEOREM 2. *Let $d = 1$. As $t \to \infty$,*

(i) $P(\eta_t(x\sqrt{t}) \neq \eta_t(y\sqrt{t})) \to 2\theta(1-\theta)\left[1 - \frac{1}{\pi}\int_0^1 \frac{e^{-|x-y|^2/4s}}{\sqrt{s(1-s)}}\,ds\right] \quad (x \neq y),$

(ii) $\{\eta_t(x\sqrt{t})\}_{x\in\mathbf{Z}} \Rightarrow \{\eta_\infty(x)\}_{x\in\mathbf{Z}},$

(iii) $t^{-1/2}\sum_{|x|\le\sqrt{t}} \eta_t(x) \Rightarrow Y.$

The limits η_∞ and Y are nontrivial, but are too complicated to allow much in the way of explicit computation. The factor $\sqrt{t}$ is the "natural scale" in the following sense. If at time t we replace $\sqrt{t}$ by $f(t)$, and if $f(t)/\sqrt{t} \to 0$ or $+\infty$, then the limits in Theorem 2 will be trivial. For instance, $\{\eta_\infty(x)\}$ will be totally correlated or independent.

There remains the case of dimension 2. Little beyond the basic clustering property (2.1) has been proved in this setting, with the notable exception of some results on occupation times [9]. One simple question to ask in the presence of clustering is whether a given site, say the origin, changes color infinitely often, or whether one of the clusters eventually surrounds the site forever. Our discussion above shows that in one dimension this problem is equivalent to whether annihilating random walks on the integers are site recurrent. This is a widely studied problem (see, e.g., [11], [14], [3]); the answer is yes, so the origin does change colors forever when $d = 1$. The same question is not quite as easy in two dimensions. Clifford and Sudbury argue for "color change i.o. at 0" when $d = 2$ in [8]; their argument does not seem rigorous to us, but is not too hard to fix. In any case, a study of occupation times yields stronger and rather surprising results. Let T_t denote the total amount of time up to time t that the origin is colored green ($= 1$). Then it turns out [9] that T_t/t converges to θ a.s. in all dimensions ≥ 2. This is not surprising for $d \geq 3$ when the model is stable; in fact Andjel and Kipnis [1] have recently obtained a pointwise ergodic theorem which applies to a large class of initial states in this case. But it is slightly paradoxical that the proportion of time a site is green converges with probability 1 even in the presence of the two dimensional clustering. (In one dimension T_t/t converges in distribution to a nonconstant limiting random variable which can be represented by means of annihilating Brownian motions.) One can also prove occupation time central limit theorems for the voter model with $d \geq 2$. Curiously, the normalization required is of the usual order $\sqrt{t}$ for $d \geq 5$, but of higher order when $d = 2$, 3, or 4. For more details and the proofs, see [9].

The discussion of the last paragraph pretty well sums up what is known about the two dimensional voter model. It is intriguing to ask whether there is a natural scale for the clustering in this case, whether there is an invariance principle anything like Arratia's, what the appropriate analogue of Theorem 2 is, and generally whether we can formulate and prove any rigorous results which illuminate the graphics of Section 1. In the next section we will try to answer these questions.

3. Results. Our first step in analysing the clustering of the two dimensional voter model is to study the correlation functions $\rho_t(A) = P(\eta_t(A) \equiv 1)$ for finite sets A. In particular, the two-point correlations give a good indication of the interdependence between two distant sites. The correlations of the voter model may be expressed in terms of an auxiliary process of *coalescing random walks* by means of a duality equation:

$$P(\eta_t(A) \equiv 1) = E\left[\theta^{\#\xi_t^A}\right] \qquad (A \text{ finite}). \tag{3.1}$$

Recall that θ is the initial density of particles for η_t. Here $\#\xi_t^A$ denotes the number of coalescing random walks still around at time t for the process started with particles on A. As the name implies, ξ_t consists of particles which undergo independent continuous time rate 1 simple random walks, except that they coalesce whenever one attempts to jump to a site occupied by another. The above duality equation has been the principal tool for studying the voter model; Theorem 1, in particular, is an easy consequence. A proof of (3.1) may be found in [23] or [15], for example.

Now suppose we take $A = \{xt^{\alpha/2}, yt^{\alpha/2}\}$ in (3.1), $0 \le \alpha \le \beta < \infty$. (To simplify notation, here and for the rest of the paper we will identify each point of $\mathbb{R}^2$ with the nearest point in $\mathbb{Z}^2$, adopting some convention in case of ties.) Since the difference of two rate 1 random walks is a rate 2 random walk, a little computation gives

$$(3.2) \qquad P\big(\eta_{t^\beta}(xt^{\alpha/2}) \neq \eta_{t^\beta}(yt^{\alpha/2})\big) = 2\theta(1-\theta)P_{(y-x)t^{\alpha/2}}(\tau > 2t^\beta),$$

where τ is the first hitting time of 0 for a rate 1 random walk. Consider first the extreme cases $\alpha = 0$ and $\alpha = \beta$. If $\alpha = 0$ then the right side of (3.2) tends to 0 as $t \to \infty$ by recurrence. This is in fact the proof of (2.1). If $\alpha = \beta$, on the other hand, the right side tends to $2\theta(1-\theta)$. (One way to get this is to use Donsker's invariance principle and the fact that Brownian motion misses points in the plane.) Already we see the contrast with dimension 1 and Theorem 2(i). What happens for intermediate power laws α? Erdös and Taylor [10] showed that the probability on the right side of (3.2) tends to α/β; the following heuristic calculation gives their answer. Use a "last time at 0" decomposition, and write $p_u(x, y) = P(\xi_u(x) = y)$ to see that

$$P_{xt^{\alpha/2}}(\tau \le 2t^\beta) = p_{2t^\beta}(0, xt^{\alpha/2}) + \int_0^{2t^\beta} p_u(0, xt^{\alpha/2})P_{(1,0)}(\tau > 2t^\beta - u)\,du.$$

Standard random walk estimates (cf. [28], pages 79 and 167) yield

$$\begin{aligned}
P_{(y-x)t^{\alpha/2}}(\tau \le 2t^\beta) &\approx \int_1^{t^\beta} \frac{e^{-|y-x|^2 t^\alpha/u}}{\pi u} \, \frac{\pi}{\log(2t^\beta - u)}\,du \\
&\approx \int_0^\beta e^{-|y-x|^2 t^{\alpha-s}} \frac{\log t}{\log(2t^\beta - t^s)}\,ds \\
&\to 1 - \alpha/\beta \qquad \text{as } t \to \infty.
\end{aligned}$$

Note that there is no dependence on $y - x$ in the limit. Thus the pair correlation in the voter model between two sites roughly distance $t^{\alpha/2}$ apart at time t converges, as $t \to \infty$, to a distinct interpolating value for each $\alpha \in (0,1)$, and these values do not depend on the actual multiple of $t^{\alpha/2}$ which separates the sites. For the closely related planar stepping stone model of population genetics, this observation is due to Sawyer [25]; his paper contains an alternate derivation of the Erdös–Taylor result.

Now it is reasonable to expect that for any n sites $x_1, x_2, \dots, x_n$ in $\mathbb{Z}^2$, $\alpha < \beta < \infty$, and $1 \le k \le n$, writing $A_t = \{x_1 t^{\alpha/2}, \dots, x_n t^{\alpha/2}\}$, one should have

$$(3.3) \qquad \lim_{t\to\infty} P\big(\#\xi_{t^\beta}^{A_t} = k\big) = p_{n,k}(\alpha/\beta),$$

for some functions $p_{n,k}$. In other words, the probability that n particles initially separated by roughly $t^{\alpha/2}$ will coalesce down to k particles at time t^β should also converge to a limit independent of the fine spatial structure. In light of (3.1), this will mean that all the limiting correlations of a $t^{\alpha/2}$ thinning of the voter model at time t will depend only on the number of sites involved, i.e., that the limiting random field is *exchangeable*. So our first major task is to compute the $p_{n,k}(\alpha)$.

As we shall see, the answer is

$$p_{n,k}(\alpha) = \sum_{j=k}^{n} a_{n,k}(j)\alpha^{\binom{j}{2}}, \tag{3.4}$$

where

$$a_{n,k}(j) = \frac{(-1)^{j+k}(2j-1)(j+k-2)!\binom{n}{j}}{k!(k-1)!(j-k)!\binom{n+j-1}{j}}.$$

The proof of (3.3)–(3.4) is rather involved, but with a little hindsight we can offer the following outline. For $n = 2$ this is simply the Erdös–Taylor result. The next case to settle is $n = k > 2$. That is to say, we wish to compute the limiting probability that no collisions occur by time t^{β} among n particles initially situated at sites separated by distance $t^{\alpha/2}$. Fix α, t, and a set A_t of n such sites. Focus on two particles indexed by i and j, say, and write

$$H^{\beta}_{\{i,j\}} = \{\text{particles } i \text{ and } j \text{ collide by } t^{\beta}\}.$$

It turns out that the complements of these events are asymptotically independent as $\{i, j\}$ ranges over all possible pairs of distinct indices, so that

$$p_{n,n}(\alpha/\beta) = \lim_{t} P\Big(\bigcap_{\{i,j\}} \tilde{H}^{\beta}_{\{i,j\}}\Big) = (\alpha/\beta)^{\binom{n}{2}}. \tag{3.5}$$

Unfortunately we do not know an intuitive argument for (3.5), so our derivation is rather devious. Here are the main steps; most details will be deferred to Section 5. Introduce:

$$F^{\beta}_{\{i,j\}} = \{\text{particles } i \text{ and } j \text{ collide first, do so by } t^{\beta}\},$$

$$q(\beta) = P(\text{some collision by } t^{\beta}).$$

Then

$$\begin{aligned} P\big(H^{\beta}_{\{i,j\}}\big) = {} & P\big(F^{\beta}_{\{i,j\}}\big) + P\big(\tilde{F}^{\beta}_{\{i,j\}}, \text{some collision by } t^{\alpha}\big) \\ & + \sum_{\{k,l\}\neq\{i,j\}} \int_{\alpha}^{\beta} P\,(\text{pair } \{k,l\} \text{ collides first, does so in } d(t^{\gamma}); \\ & \qquad \text{pair } \{i,j\} \text{ hits in } (t^{\gamma}, t^{\beta}]\,i). \end{aligned} \tag{3.6}$$

We will show in Section 5 that asymptotically the integrand on the right factors as

$$P(\{k,l\} \text{ pair hits first, in } d(t^{\gamma}))\, P(\{i,j\} \text{ pair hits in } (t^{\gamma}, t^{\beta}]).$$

The second term here is $\approx 1 - \gamma/\beta$ since the particles are separated by about $t^{\gamma/2}$ at time t^{γ}. Integrating by parts we get

$$P\big(H^{\beta}_{\{i,j\}}\big) \approx P\big(F^{\beta}_{\{i,j\}}\big) + \frac{1}{\beta} \sum_{\{k,l\}\neq\{i,j\}} \int_{\alpha}^{\beta} P\big(F^{\gamma}_{\{k,l\}}\big)\, d\gamma.$$

The left side is $\approx 1 - \alpha/\beta$; sum over all $\{i, j\}$ to find that

$$\binom{n}{2}(1 - \alpha/\beta) = q(\beta) + \frac{1}{\beta}\left[\binom{n}{2} - 1\right]\int_\alpha^\beta q(\gamma)\, d\gamma.$$

The solution is $q(\beta) = 1 - (\alpha/\beta)^{\binom{n}{2}}$, so (3.5) should hold in the limit as $t \to \infty$.

Now we can derive (3.3)–(3.4) for general n and k by induction. The heuristic is much the same as above: (3.5) gives the distribution of the first hiting time σ for some pair of particles, and at this time the remaining $n - 1$ particles are spaced about $t^{\gamma/2}$ apart. These must in turn coalesce down to k particles in time $t^\beta - t^\gamma \sim t^\beta$. So

$$P\left(\#\xi_{t^\beta}^{A_t} = k\right) \approx \int_\alpha^\beta P(\sigma \in d(t^\gamma))P\left(\#\xi_{t^\beta}^{A_t} = k \middle| \#\xi_{t^\gamma}^{A_t} = n - 1\right)$$

$$\approx \int_\alpha^\beta \binom{n}{2}\alpha^{\binom{n}{2}}\gamma^{-\binom{n}{2}-1} p_{n-1,k}(\gamma/\beta)\, d\gamma.$$

A routine calculation verifies that the right side equals $p_{n,k}(\alpha/\beta)$ as given in (3.4). Let us pause here to summarize the findings so far, by stating a precise version of our first major result. Some uniformity is needed to make the induction rigorous; look at Section 5 for the details of the proof.

THEOREM 3. *Let ξ_t^A denote the system of coalescing random walks starting from $A \subset \mathbb{Z}^2$. Fix $0 < \alpha_0 \le \beta_0 < \infty$ and $0 < c < \infty$. Then for each n, uniformly in α, β and $A = \{x_1, \dots, x_n\}$ such that $\alpha_0 \le \alpha \le \beta \le \beta_0$ and*

$$\frac{t^{\alpha/2}}{c \log t} \le \|x_i - x_j\| \le c t^{\alpha/2} \log t \qquad (i \ne j),$$

we have

$$\lim_{t\to\infty} P\left(\#\xi_{t^\beta}^{A} = k\right) = p_{n,k}(\alpha/\beta) \qquad (1 \le k \le n),$$

where the $p_{n,k}$ are given by (3.4).

As we pointed out earlier, since the correlation functions are distribution determining, (3.1) and (3.3) together imply that the $t^{\alpha/2}$ thinning of the voter model at time t converges to a limiting field for each $\alpha \in [0, 1]$, i.e.,

$$\{\eta_t^\theta(xt^{\alpha/2})\}_{x\in\mathbb{Z}^2} \Rightarrow \{\eta_\infty^{\theta,\alpha}(x)\}_{x\in\mathbb{Z}^2} \quad \text{as } t \to \infty. \tag{3.7}$$

The absence of spatial structure in the limiting correlations means that the limit field is exchangeable, and hence de Finetti's theorem asserts that this field is a mixture $dF_{\theta,\alpha}(s)$ of Bernoulli product measures with density s. In particular,

$$\rho_\infty^{\theta,\alpha}(A) = P\left(\eta_\infty^{\theta,\alpha}(A) \equiv 1\right) = \int_0^1 s^n\, dF_{\theta,\alpha}(s), \qquad A = \{x_1, \dots, x_n\}. \tag{3.8}$$

The next order of business is to determine the mixture $dF_{\theta,\alpha}$. To do so we differentiate the integral formula

$$p_{n,k}(\alpha) = \int_\alpha^1 \binom{n}{2} \alpha^{\binom{n}{2}} \gamma^{-\binom{n}{2}-1} p_{n-1,k}(\gamma)\,d\gamma \tag{3.9}$$

and make the change of variables $\alpha = e^{-u}$ to get the system of differential equations

$$\frac{d}{du} p_{n,k}(e^{-u}) = -\binom{n}{2} p_{n,k}(e^{-u}) + \binom{n}{2} p_{n-1,k}(e^{-u}),$$
$$p_{n,k}(e^{-u})|_{u=0} = 1_{\{n=k\}}.$$

These are precisely the backward equations for the transition function of a pure death process D_t on the positive integers which jumps from n to $n-1$ at exponential rate $\binom{n}{2}$ $(n \geq 2)$. Hence

$$p_{n,k}(\alpha) = P_n(D_{\log(1/\alpha)} = k). \tag{3.10}$$

So now we have a connection (3.1) between the voter model η and the coalescing random walks ξ, and a connection (3.10) between ξ and the death process D. One more link is needed to discover the mixture dF; our mystery guest in the scenario for critical clustering turns out to be the *Fisher–Wright diffusion* Y_t. Recall that Y_t is the strong Markov process on $[0,1]$ with generator

$$Gf(\gamma) = \tfrac{1}{2}\gamma(1-\gamma)f''(\gamma) \qquad 0 < \gamma < 1;$$

let $q_t(\gamma, \cdot) = P_\gamma(Y_t \in \cdot)$ denote the corresponding transition function. It is well-known in the mathematical genetics literature that Y_t and D_t satisfy the duality equation

$$E_\theta[Y_t^n] = E_n[\theta^{D_t}] \qquad n \geq 1, 0 \leq \theta \leq 1. \tag{3.11}$$

Note the similarity to (3.1). Tavaré [29] has a nice survey of genetics problems related to ours, and discusses identity (3.11). (Actually there is a whole family of Fisher–Wright diffusions. Geneticists would label ours the one with *pure drift*, whereas probabilists would call it the one with *no drift*. To avoid confusion, we just call Y_t *the* Fisher–Wright diffusion.)

Putting together the pieces, we can finally compute the desired de Finetti mixture in just three lines:

$$\begin{aligned}
\rho_\infty^{\theta,\alpha}(A) &= \sum_{k=1}^{\infty} \theta^k p_{n,k}(\alpha) && \text{by (3.1) and (3.3)},\\
&= E_n[\theta^{D_{\log(1/\alpha)}}] && \text{by (3.10)},\\
&= \int_0^1 s^n q_{\log(1/\alpha)}(\theta, ds) && \text{by (3.11)}.
\end{aligned}$$

Comparing with (3.8), we see that the mixture $F_{\theta,\alpha}$ is simply the distribution of the Fisher–Wright diffusion starting from θ at time $\log(1/\alpha)$. This, then, is our second major result.

THEOREM 4. *The $t^{\alpha/2}$ thinning of the two dimensional voter model at time t converges as $t \to \infty$ to a limiting exchangeable random field with de Finetti mixture $F_{\theta,\alpha} = P_\theta(Y_{\log(1/\alpha)} \in \cdot)$, where Y is the Fisher–Wright diffusion. In other words,* (3.7)–(3.8) *hold for this F.*

Theorem 4 is extremely suggestive of a more elaborate result; namely, that the limiting density of the $t^{\alpha/2}$ thinning diffuses as a process in α. One can make this precise, but we prefer to take a somewhat different tack in order to address our central question of cluster size. Namely, introduce the block averages

$$B^t(\alpha) = \frac{1}{4t^\alpha} \sum_{\|x\| \le t^{\alpha/2}} \eta_t(x).$$

Since the overwhelming proportion of sites in the block $\{\|x\| \le t^{\alpha/2}\}$ are order $t^{\alpha/2}$ apart, it is easy to conclude from Theorem 4 that as $t \to \infty$, $B^t(\alpha) \Rightarrow Y_{\log(1/\alpha)}$ (α fixed). Then one suspects that

$$B^t(\cdot) \overset{t\to\infty}{\Rightarrow} Y_{\log(1/\cdot)} \quad \text{(as processes)}. \tag{3.12}$$

This is indeed the case, and constitutes the main result of the paper. Using it one can deduce quite a bit about the clustering of η_t, as we shall see in the next section. Let us conclude this section by stating the precise formulation of (3.12) which we will prove, and by suggesting its qualitative meaning. For the purposes of the present paper, (3.12) simply asserts convergence of finite dimensional distributions; this turns out to be enough for the main application we have in mind. There is little doubt that (3.12) actually holds in the sense of weak convergence on path space, but the methods needed to prove tightness will be quite different from the ones featured here, so we choose to defer this question for future study. To recap, we will prove:

THEOREM 5. *As $t \to \infty$, the density of* 1's (= *green sites*) *in the two dimensional voter model on a box of side $2t^{\alpha/2}$ converges as a process in $\alpha \in (0,1]$ to a time change of the Fisher–Wright diffusion. Specifically,* (3.12) *holds with convergence in the sense of finite dimensional distributions.*

The proof of Theorem 5 depends on a generalization of Theorem 3, in which coalescing walks are allowed to start spread out by different power laws. One uses (3.1) and the generalization (Theorem 6) to estimate mixed moments of different power law block averages. This is all rather technical, so we will postpone the argument until Section 5.

A few words about the intuitive meaning of Theorem 5 are helpful at this point. Fix a very large time t, and think of the block process $B^t(\alpha)$ as running back from $\alpha = 1$ to $\alpha = 0$. Near $\alpha = 1$ most of the sites in the box are almost $t^{1/2}$ apart, hence essentially uncorrelated, and so the block average will be almost deterministically the global density θ. As α decreases the block average *diffuses*, reflecting a very strong correlation between the averages on boxes with two nearby power laws. Since 0 and 1 are accessible traps for Y_t, at some random time this diffusion is absorbed at the boundary of $[0, 1]$, reflecting the fact that the box has been swallowed by a cluster of green or black sites. This can occur at any intermediate power law α, so there is an enormous variation in the sizes of the clusters. Note that $\alpha = 0$ corresponds to time ∞ for Y, by which point the box has certainly been swallowed.

4. Applications. In this section we briefly describe various applications and extensions of our theorems. More detailed investigations with proofs will appear elsewhere.

To apply our main result, Theorem 5, let us suppose that it holds in the stronger sense of an invariance principle, i.e., convergence of processes takes place. As already mentioned, we hope to establish tightness in a subsequent paper. By applying various functionals to the block sums $B^t(\alpha)$, one should be able to read off asymptotic statistics which quantify the clustering of the two dimensional voter model. A couple of interesting examples easily come to mind. As a warm-up, take

$$M^t = \sup\{B^t(\alpha) \colon 0 < \alpha \le 1\}.$$

Then Theorem 5 suggests that M^t converges in distribution as $t \to \infty$ to the maximum value of the diffusion $Y_{\log(1/\cdot)}$ over all time. But Y has the same hitting probabilities as Brownian motion with absorption at 0 and 1, so one easily deduces that for $x \in (\theta, 1)$,

$$P(M^t \le x) \to 1 - \frac{\theta}{x} \quad \text{as } t \to \infty.$$

In this manner we determine the asymptotic distribution of the largest density of green color on any block centered at the origin. (It is easy to see that blocks of side bigger than $\sqrt{t}$ will have asymptotic density θ.)

Now for our main application. Introduce the functional

$$L^t = \sup\{\alpha > 0 \colon B^t(\alpha) = 0 \text{ or } 1\}.$$

Evidently, L^t measures the power law of the largest box centered at the origin which is all one color at time t. The invariance principle heuristic suggests that as $t \to \infty$, L^t should converge in distribution to the time when $Y_{\log(1/\cdot)}$ first hits the boundary at 0 or 1. The distribution of this hitting time is readily computed using (3.10) and (3.11). (Alternatively, consult [13] or [29].) We are lead to suspect that for $\gamma \in (0, 1]$,

$$P(L^t \le \gamma) \to F_\theta(\gamma) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} \left(\theta^k + (1 - \theta)^k\right) p_{\infty, k}(\gamma) \tag{4.1}$$

as $t \to \infty$, with

$$p_{\infty, k}(\gamma) = \lim_{n \to \infty} p_{n, k}(\gamma) = \sum_{j=k}^{\infty} \frac{(-1)^{j+k}(2j - 1)(j + k - 2)!}{k!(k - 1)!(j - k)!} \gamma^{\binom{j}{2}}.$$

Rewording things slightly, let N^t denote the width of the largest box centered at the origin which is a solid color at time t. Our result is that

$$\frac{\log N^t}{\log t} \Rightarrow L^\infty/2 \quad \text{as } t \to \infty, \tag{4.2}$$

where L^∞ has distribution function F_θ. In words, (4.2) asserts that the cluster containing the origin at time t has a size which is randomly distributed over the powers of t between 0 and $\frac{1}{2}$, and that the distribution of this random power obeys a hypergeometric law prescribed by (4.1). Of course size is a rather elusive notion here; anyone familiar with the most basic results on percolation in the plane (cf. [19]) will recall that already at time 0, for θ near 0 or 1 the connected component which shares the color of the origin is infinite with positive probability. Rather than confront the intricacies of connectivity, we measure the size of a cluster by how large a box will fit inside of it. With this proviso, (4.2) is a rather precise mathematical formulation of the dynamic suggested by the pictures in Section 1.

Actually, one does not need an invariance principle, or even the full strength of Theorem 5, to prove (4.2). Observe that

$$P\left(\frac{\log N^t}{\log t} \le \frac{\gamma}{2}\right) = P\big(B^t(\gamma) = 0 \text{ or } 1\big),$$

i.e., only a marginal limit distribution is involved. On the other hand, 0 and 1 are not continuity points of the Fisher–Wright diffusion at any positive time, so our main application is *not* an immediate consequence of Theorem 5. One must check that mass near the boundary at finite times is not lost in the limit. This amounts to showing that if a block average at time t on a box of size t^γ equals $\delta \approx 0$, then a box of size $t^{\gamma-\varepsilon}$ must have *no* green sites with overwhelming probability. The needed patch turns out to take nearly as much effort as Theorem 5 itself, but we have obtained a rigorous proof of (4.2) in joint work with Maury Bramson.

A brief digression is in order here to explain the nature of the patch. One basic question about the voter model which we did not mention in Section 2 concerns the behavior of the process η_t^0 starting with a single green pixel surrounded by a sea of black. (In terms of donkeys and elephants, think of the Minnesota electorate in the '84 presidential election.) It is easy to see that the number of green cells evolves as a time change of simple random walk with a positive minimal jump rate on the positive integers and absorption at 0. So the survival probability

$$p_t = P\big(\eta_t^0(x) \not\equiv 0\big)$$

clearly tends to 0, and one tries to find the asymptotics. In [6] it was shown that

$$\begin{aligned} p_t &\sim (\pi t)^{-1/2}, & d &= 1, \\ &\sim \log t/\pi t, & d &= 2, \\ &\sim (\gamma_d t)^{-1}, & d &\ge 3, \end{aligned} \tag{4.3}$$

where γ_d is the probability that simple random walk on $\mathbb{Z}^d$ never returns to its initial position. The derivation of (4.3) consists of a tightness patch to a theorem of Sawyer [26] concerning the "stepping stone model" which will be mentioned below. A similar but more intricate analysis is necessary to pass from Theorem 5 to (4.2); the argument will appear in a forthcoming paper [7].

Our principal focus in this study is the voter model, so we are inclined to view the coalescing random walks ξ_t as an auxiliary process. But variants on our techniques yield some results for coalescing walks which are interesting in their own right. For instance, suppose we start with particles at the sites of $\sqrt{t}\,\mathbb{Z}^2$, and observe the system at times t^γ, $\gamma \geq 1$. Recalling Theorem 4, it is natural to expect convergence to a limiting process in γ. Convergence does indeed take place (in the sense of finite dimensional distributions, and presumably on path space). By keeping track of which particles collide with which others, we are able to show that the limit is a simple time change of Kingman's coalescent [22]. Again, the details will appear elsewhere. For now, let us simply mention a colorful result concerning ξ_t which makes use of our joint work with Maury Bramson [7]. Namely, suppose one starts with particles on a *solid* block of side $2t^{\alpha/2}$, and asks for the probability P_t that they *all* have coalesced down to a single particle by time t. Then as $t \to \infty$,

$$
\begin{aligned}
P_t \to p_{\infty,1}(\alpha) &= 1 - 3\alpha + 5\alpha^3 - 7\alpha^6 + 9\alpha^{10} - 11\alpha^{15} + \cdots \\
&= \left[(1-\alpha)(1-\alpha^2)(1-\alpha^3)(1-\alpha^4)\cdot\cdots\right]^3.
\end{aligned}
\tag{4.4}
$$

The last amazing equality is due to Jacobi [18].

Another generalization of our results deals with more than two colors (= states). For each $N = 2, 3, \ldots$ there is an N-color voter model in which the color at any site is replaced at rate 1 by whatever color occupies a random neighboring site. There is even a model with $N = \infty$, where every site of the lattice starts off with its own distinct color. These many-colored processes are simplified versions of systems known as "stepping stone models" in the mathematical genetics literature (see e.g., [13], [20], [21], [24], [25], [26], [29]). Indeed, much of the general theory for the voter model which we described in Section 2 was anticipated by the mathematical geneticists. For instance, versions of the duality equation (3.1) are implicit already in [24] and [21], and one can argue that [24] foreshadows Theorem 1. Of this literature, Sawyer's paper [25] comes closest to our work. In Theorem 3 display (4) of that paper he shows that the two point correlations of the planar stepping stone model are spread out over the powers. He talks of a "wave of advance," and notes that one must plot on a logarithmic scale to "obtain a limiting distribution." Our (4.2) is one rigorous mathematical expression of Sawyer's observation. We plan to discuss the implications of our work for the stepping stone model in another article. For now, suffice it to say that in two dimensions there are quantities like our $p_{n,k}(\alpha)$ which quantify N-color clustering; and that for large t, the proportions of sites occupied by the various colors on blocks of side $2t^\alpha$ diffuse after the fashion of Theorem 5. The limiting diffusion, which lives on $\{x \in \mathbb{R}^N\colon \sum x_i = 1\}$, is a time change of one obtained by Ethier and Kurtz in [12]. In the case $N = \infty$ the diffusion instantaneously jumps down to a finite dimensional subspace; (4.4) gives the limiting probability that a box of side $2t^{\alpha/2}$ is all one color at time t. More details will be forthcoming.

Finally, a few words about the "domain of attraction" of our results. Although we have chosen to consider only Bernoulli product measures as initial states, our limit theorems undoubtedly hold for a large class of stationary ergodic initial distributions. Moreover, it is not hard to check that precisely the same asymptotics apply to any of the Holley–Liggett voter models [16] determined by a translation invariant transition density $p(x)$ on $\mathbb{Z}^2$ which is irreducible and has bounded range (or enough moments). There is nothing special about the box norm; any equivalent norm gives identical results. For these reasons we suspect that the critical clustering which takes place in the two dimensional voter model may be representative of more general patterns of planar clustering in homogeneous random media.

5. Proofs. This section contains rigorous versions of the heuristic arguments presented in Section 3. We have four main tasks:

A. Formulate an appropriate version of the Erdös–Taylor result;
B. Establish (3.5);
C. Prove Theorem 3;
D. Prove Theorem 5.

To begin, we introduce a little more notation. For the rest of the paper we fix α_0, β_0, and c such that

$$0 < \alpha_0 \le \beta_0 < \infty \quad \text{and} \quad 1 < c < \infty.$$

Let C denote a finite absolute constant whose value may change from line to line; and let $\varepsilon(t; a, b, \dots)$ be a function of t, depending only on $a, b, \dots$, which tends to 0 as $t \to \infty$. Uniformity will be needed to carry out the numerous estimates below, so we will let sites x of magnitude $t^{\alpha/2}$ have norms in the larger interval

$$\Gamma_t(c, \alpha) = \left[\frac{t^{\alpha/2}}{c \log t}, ct^{\alpha/2} \log t\right].$$

A. *Erdös–Taylor.* Recall that τ is the first time a simple random walk hits the origin. Our first preliminary result is as follows.

PROPOSITION 1.

(i) $$\lim_{t\to\infty} \sup_{\substack{\alpha_0 \le \alpha \le \beta \le \beta_0,\ t^\alpha \le s \le t^\beta \\ \|x\| \in \Gamma_s(c,1)}} \left| P_x(\tau \le s) - \left(1 - \frac{2\log\|x\|}{\log s}\right)\right| = 0$$

and

(ii) $$\lim_{t\to\infty} \sup_{\substack{\alpha_0 \le \alpha \le \beta \le \beta_0,\ t^\alpha \le s \le t^\beta \\ \|x\| \in \Gamma_s(c,1)}} \left| P_x(\tau \le t^\beta - s) - \left(1 - \frac{2\log\|x\|}{\beta \log t}\right)\right| = 0.$$

In [10], Erdös and Taylor prove a somewhat stronger version of (i) for discrete time simple random walk. The relevant equations are (2.16)–(2.18) of that paper. The translation to our continuous time setting is straightforward but tedious, so we omit it. (ii) can be verified using the same techniques.

B. *Argument for* (3.5). We will now prove:

PROPOSITION 2. *Uniformly in* α, β *and* $A = (x_1, \dots, x_n)$ *such that* $\alpha_0 \le \alpha \le \beta \le \beta_0$ *and* $\|x_i - x_j\| \in \Gamma_t(c, \alpha)$ *for all* $i \ne j$,

$$\lim_{t\to\infty} P\left(\#\xi^A_{t^\beta} = n\right) = 1 - \left(\frac{\alpha}{\beta}\right)^{\binom{n}{2}}.$$

As we have already warned the reader, Proposition 2 takes some work. Our proof requires two lemmas. The first one is technical, and deals with joint distributions of random walks. The second formalizes our derivation of the integral equation for $q(\beta)$ in Section 3. We will first state the lemmas, next show how they prove Proposition 2, and finish by proving the lemmas.

In what follows, let $X_s(x)$ $(s \ge 0,\ x \in \mathbb{Z}^2)$ be independent continuous time rate 1 simple random walks starting from each site $x \in \mathbb{Z}^2$, $\tau_{i,j}$ the first hitting time of the walks from x_i and x_j. As before, σ will be the first collision time among walks from A.

LEMMA 1. *Uniformly for* $\alpha \in (\alpha_0, \infty)$ *and* $\|x_i - x_j\| \in \Gamma_t(c, \alpha)$ $(i \ne j)$,

$$\text{(i)}\quad \lim_{t\to\infty} \int_{t^\alpha}^{\infty} P\left(X_s(x_1) = X_s(x_2),\ \|X_s(x_1) - X_s(x_3)\| \notin \Gamma_s(4c, 1)\right) ds = 0$$

and

$$\text{(ii)}\quad \lim_{t\to\infty} \int_{t^\alpha}^{\infty} P\left(X_s(x_1) = X_s(x_2),\ \|X_s(x_3) - X_s(x_4)\| \notin \Gamma_s(4c, 1)\right) ds = 0.$$

LEMMA 2. *Suppose that for each* $t > 0$, $q_t(\alpha, \beta)$ *is measurable and for* $\alpha_0 \le \alpha \le \beta \le \beta_0$,

$$q_t(\alpha, \beta) = \binom{n}{2}\left[1 - \frac{\alpha}{\beta}\right] - \frac{1}{\beta}\left[\binom{n}{2} - 1\right]\int_\alpha^\beta q_t(\alpha, \gamma)\, d\tau + \varepsilon(t, \alpha, \beta),$$

where

$$\lim_{t\to\infty} \sup_{\alpha_0 \le \alpha \le \beta \le \beta_0} |\varepsilon(t, \alpha, \beta)| = 0.$$

Then as $t \to \infty$,

$$q_t(\alpha, \beta) \to 1 - \left(\frac{\alpha}{\beta}\right)^{\binom{n}{2}} \quad \textit{uniformly in } \alpha_0 \le \alpha \le \beta \le \beta_0.$$

Let us see how the lemmas yield Proposition 2. Our starting point is (3.6). The summands on the right side are of two types: $\{k, l\} \cap \{i, j\} \ne \varnothing$ and $\{k, l\} \cap \{i, j\} = \varnothing$. By Proposition 1 and the Markov property a typical term of the first type can be written as

$$\varepsilon(t;\alpha,\beta,A) + \int_{t^\alpha}^{t^\beta}\Bigg[\sum_{y,\,z:\ \|y-z\|\notin\Gamma_s(4c,1)} + \sum_{y,\,z:\ \|y-z\|\in\Gamma_s(4c,1)}\Bigg] \times P\big(\sigma=\tau_{i,k}\in ds,\ \xi_s^{x_i}=y,\ \xi_s^{x_j}=z\big)P\big(\xi_{t^\beta-s}^{y}=\xi_{t^\beta-s}^{z}\big),$$

where $|\varepsilon(t;\alpha,\beta,A)|\le\varepsilon(t;\alpha_0,c)$. As $t\to\infty$, the part of this last integral contributed by the first sum tends to 0 uniformly in α and A by Lemma 1(i). By Proposition 1 the second portion is

$$\int_{t^\alpha}^{t^\beta}\left(1-\frac{\log s}{\beta\log t}\right)P(\sigma=\tau_{i,k}\in ds)+\varepsilon(t;\alpha,\beta,A,c),$$

where $\varepsilon(t;\alpha,\beta,A,c)\le\varepsilon(t;\alpha_0,\beta_0,c)$. Integration by parts and a change of variables yields

$$\int_{t^\alpha}^{t^\beta}P\big(\sigma=\tau_{i,k}\in ds,\ s<\tau_{i,j}\le t^\beta\big)=\frac{1}{\beta}\int_\alpha^\beta P\big(F_{\{i,k\}}^{\gamma}\big)\,d\gamma+\varepsilon(t;\alpha,\beta,A,c).$$

Using Lemma 2(ii) for terms of the second type, it follows that

$$P\big(H_{\{i,j\}}^{\beta}\big)=P\big(F_{\{i,j\}}^{\beta}\big)+\frac{1}{\beta}\sum_{\{k,l\}\ne\{i,j\}}\int_\alpha^\beta P\big(F_{\{k,l\}}^{\gamma}\big)\,d\gamma+\varepsilon(t;\alpha,\beta,A,c).$$

By Proposition 1 again the left side is $1-(\alpha/\beta)+\varepsilon(t;\alpha,\beta,A)$. Sum over all pairs and apply Lemma 2 to get the desired result. □

PROOF OF LEMMA 1. Our strategy is to decompose the probabilities appropriately and appeal to standard random walk estimates ([24]). To prove part (i) we split the integral into four pieces, denoted I_1, I_2, I_3, and I_4, by considering separately sample paths where

$$(I_1)\qquad \left\{\|X_s(x_1)\|<\frac{s^{1/2}}{2c\log s}\right\},$$

$$(I_2)\qquad \left\{\|X_s(x_1)\|>2cs^{1/2}\log s\right\},$$

$$(I_3)\qquad \left\{\|X_s(x_1)\|\in\Gamma_s(2c,1),\ \|X_s(x_1)-X_s(x_3)\|<\frac{s^{1/2}}{4c\log s}\right\},$$

$$(I_4)\qquad \left\{\|X_s(x_1)\|\in\Gamma_s(2c,1),\ \|X_s(x_1)-X_s(x_3)\|>4cs^{1/2}\log s\right\},$$

respectively. The four integrals are estimated as follows:

$$I_1\le C\int_{t^{\alpha_0}}^{\infty}\#\left\{\|y\|<\frac{s^{1/2}}{2c\log s}\right\}\frac{1}{s^2}\,ds,$$

$$I_2\le C\int_{t^{\alpha_0}}^{\infty}\sum_{\|y\|>cs^{1/2}\log s}p_s(x_1,y)\frac{1}{s}\,ds$$

$$\le C\int_{t^{\alpha_0}}^{\infty}\frac{1}{s\log^2 s}\,ds.$$

(On this piece we use the fact that $\|x_1 - X_s(x_1)\| \geq \|X_s(x_1)\|/2$ and Chebyshev's inequality.)

$$I_3 \leq C\int_{t^{\alpha_0}}^{\infty} P(X_s(x_1) \in \Gamma_s(2c,1))\#\left\{\|y\| < \frac{s^{1/2}}{4c\log s}\right\}\frac{1}{s^2}\,ds$$

$$\leq C\int_{t^{\alpha_0}}^{\infty} \frac{1}{s\log^2 s}\,ds,$$

$$I_4 \leq C\int_{t^{\alpha_0}}^{\infty} P(\|X_s(x_3)\| > 2cs^{1/2}\log s)\frac{1}{s}\,ds \leq C\int_{t^{\alpha_0}}^{\infty}\frac{1}{s\log^2 s}\,ds.$$

(Here we use $\|X_s(x_3)\| > 2cs^{1/2}\log s$, and Chebyshev again.) It is easy to check that each of the above bounds is $\varepsilon(t;\,\alpha_0, c)$, which proves part (i). The proof of (ii) is similar, but easier because there is less dependence. Rewrite the integral as

$$\int_{t^{\alpha_0}}^{\infty} P(X_{2s}(x_1 - x_2) = 0)P(\|X_{2s}(x_3 - x_4)\| \notin \Gamma_s(4c,1))\,ds.$$

Estimating as above, we majorize the integral by

$$C\int_{t^{\alpha_0}}^{\infty}\frac{1}{s}\left[\frac{s}{\log^2 s}\frac{1}{s} + \frac{1}{\log^2 s}\right]ds = \varepsilon(t;\,\alpha_0, c),$$

as desired. □

PROOF OF LEMMA 2. Write:

$$q(\alpha,\beta) = 1 - \left(\frac{\alpha}{\beta}\right)^{\binom{n}{2}},$$

$$\delta_t(\alpha,\beta) = q_t(\alpha,\beta) - q(\alpha,\beta),$$

$$\bar{\delta}_t(\alpha,\beta) = \sup_{\alpha\leq\bar{\alpha}\leq\bar{\beta}\leq\beta} |\delta_t(\bar{\alpha},\bar{\beta})|\,.$$

Since q satisfies our integral equation with equality,

$$|\delta_t(\alpha,\beta)| \leq \left[\binom{n}{2} - 1\right]\frac{1}{\beta}\int_{\alpha}^{\beta}\bar{\delta}_t(\alpha,\gamma)\,d\gamma + \varepsilon(t;\,\alpha_0,\beta_0)$$

$$\leq \frac{1}{2}\bar{\delta}_t(\alpha_0, r\alpha_0) + \varepsilon(t;\,\alpha_0,\beta_0) \qquad (\alpha_0 \leq \alpha \leq \beta \leq r\alpha_0),$$

where $r > 1$ is chosen so that $[\binom{n}{2} - 1](1 - r^{-1}) = \frac{1}{2}$. Hence

$$\bar{\delta}_t(\alpha_0, r\alpha_0) \leq 2\varepsilon(t;\,\alpha_0,\beta_0),$$

i.e., $g_t(\alpha,\beta) \to g(\alpha,\beta)$ as $t\to\infty$ on $(\alpha_0, r\alpha_0)$. Now iterate the argument k times, where $r^k\alpha_0 \geq \beta_0$, to finish the proof. □

C. *Proof of Theorem 3.* As suggested in Section 3, the proof of Theorem 3 proceeds by induction. The induction hypothesis is

$$\lim_{t\to\infty} P(\#\xi_{t^\beta}^A = k) = p_{n,k}(\alpha/\beta) \quad \text{for } 1 \leq k \leq n,$$

uniformly in α, β and $A = \{x_1, \dots, x_n\}$ such that $\alpha_0 \le \alpha \le \beta \le \beta_0$ and $\|x_i - x_j\| \in \Gamma_t(c, \alpha)$ for all $i \neq j$. If $n = 2$ this follows immediately from Proposition 2. The induction step is to prove

$$(5.1) \qquad \lim_{t \to \infty} P\left(\#\xi^B_{t^\beta} = k\right) = \binom{n+1}{2} \alpha^{\binom{n+1}{2}} \int_\alpha^\beta \gamma^{-\binom{n+1}{2}-1} p_{n,k}(\gamma/\beta)\, d\gamma$$

uniformly in α, β and $B = \{x_1, \dots, x_{n+1}\}$ such that $\alpha_0 \le \alpha \le \beta \le \beta_0$, $\|x_i - x_j\| \in \Gamma_t(c, \alpha)$ for all $i \neq j$. The probability that any pair of walks collides by t^α tends uniformly to 0 as $t \to \infty$, so

$$(5.2) \qquad \begin{aligned} P\left(\#\xi^B_{t^\beta} = k\right) &= \int_{t^\alpha}^{t^\beta} P\left(\sigma \in ds, \#\xi^B_{t^\beta} = k\right) + \varepsilon(t; B, \beta) \\ &= \int_{t^\alpha}^{t^\beta} \sum_{A = \{y_1, \dots, y_n\}} P\left(\sigma \in ds, \xi^B_s = A\right) P\left(\#\xi^A_{t^\beta - s} = k\right) + \varepsilon(t; B, \beta), \end{aligned}$$

where $|\varepsilon(t; B, \beta)| \le \varepsilon(t; \alpha_0, \beta_0, c)$. To decompose the last integral we need estimates which generalize Proposition 1 and Lemma 1.

PROPOSITION 3. *For all α, β, A, and s such that $\alpha_0 \le \alpha \le \beta \le \beta_0$, $t^\alpha \le s \le t^\beta$, and $A = \{x_1, \dots, x_n\}$ with $\|x_i - x_j\| \in \Gamma_s(c, 1)$ for all $i \neq j$,*

$$P\left(\#\xi^A_{t^\beta - s} = k\right) = p_{n,k}\left(\frac{\log s}{\beta \log t}\right) + \varepsilon(t; A, \beta, s),$$

where $|\varepsilon(t; A, \beta, s)| \le \varepsilon(t; \alpha_0, \beta_0, c)$.

PROOF.

$$\begin{aligned} &\left|P\left(\xi^A_{t^\beta} = k\right) - P\left(\xi^A_{t^\beta - s} = k\right)\right| \\ &\quad \le 2P\left(\xi^A_\cdot \text{ has a collision in } [t^\beta - s, t^\beta]\right) \\ &\quad \le 2 \sum_{\{i,j\}} P\left(t^\beta - s \le \tau_{i,j} \le t^\beta\right) \le \varepsilon(t; \alpha_0, \beta_0, c) \end{aligned}$$

by Proposition 1. So the claim follows by induction. □

LEMMA 3. *Let $G_s = \{\|y - z\| \in \Gamma_s(4c, 1)$ for all distinct $y, z \in \xi^A_s\}$. Then*

$$\int_{t^\alpha}^{t^\beta} P(\sigma \in ds, \tilde{G}_s) \le \varepsilon(t; \alpha_0, \beta_0, c).$$

PROOF. The integral is majorized by

$$\begin{aligned} &\sum_{\{i,j\}} \sum_{\{k,l\} \neq \{i,j\}} \int_{t^\alpha}^\infty P\left(\sigma = \tau_{k,l} \in ds, \|\xi^{x_i}_s - \xi^{x_j}_s\| \notin \Gamma_s(4c, 1)\right) \\ &\quad \le \sum_{\{i,j\}} \sum_{\{k,l\} \neq \{i,j\}} \int_{t^\alpha}^\infty P\left(X_s(x_k) = X_s(x_l), \|X_s(x_i) - X_s(x_j)\| \notin \Gamma_s(4c, 1)\right) ds \\ &\quad \le \varepsilon(t; \alpha_0, c) \quad \text{by Lemma 1.} \ \square \end{aligned}$$

Now using Proposition 2 and Lemma 3, we can rewrite the integral in (5.2) as

$$\int_{t^\alpha}^{t^\beta} P(\sigma \in ds, G_s)\, p_{n,k}\left(\frac{\log s}{\beta \log t}\right) + \varepsilon(t; A, \alpha, \beta)$$
$$= \int_{t^\alpha}^{t^\beta} P(\sigma \in ds)\, p_{n,k}\left(\frac{\log s}{\beta \log t}\right) + \varepsilon(t; A, \alpha, \beta).$$

Finally, we apply Proposition 2 and integrate by parts twice to get (5.1), thereby completing the induction step.

D. *Proof of Theorem 5.* The first step is to formulate and prove a generalization of Theorem 3 to coalescing walks starting from sites on different power scales. To state the result we need a little more notation. For $k \geq 1$, positive integers $n_1, n_2, \ldots, n_k$ and m, and $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_k < \beta < \infty$, put $p_{n_1; m}(\alpha_1; \beta) = p_{n_1, m}(\alpha_1/\beta)$ and set

$$(5.3)\qquad \begin{aligned} &p_{n_1, \ldots, n_k; m}(\alpha_1, \ldots, \alpha_k; \beta) \\ &\quad = \sum_{i_1, \ldots, i_{k-1}} p_{n_1, i_1}(\alpha_1/\alpha_2) p_{n_2 + i_1, i_2}(\alpha_2/\alpha_3) \cdot \cdots \\ &\qquad\qquad \times p_{n_{k-1} + i_{k-2}, i_{k-1}}(\alpha_{k-1}/\alpha_k) p_{n_k + i_{k-1}, m}(\alpha_k/\beta). \end{aligned}$$

It is easy to see from (5.3) that

$$(5.4)\qquad \begin{aligned} &p_{n_1, \ldots, n_k; m}(\alpha_1, \ldots, \alpha_k; \beta) \\ &\quad = \sum_l p_{n_1, \ldots, n_{k-1}; l}(\alpha_1, \ldots, \alpha_{k-1}; \alpha_k) p_{n_k + l, m}(\alpha_k/\beta) \\ &\quad = \sum_l p_{n_1, l}(\alpha_1/\alpha_2) p_{n_2 + l, n_3, \ldots, n_k; m}(\alpha_2, \ldots, \alpha_k; \beta). \end{aligned}$$

THEOREM 6. *Fix $0 < \alpha_0 \leq \beta_0 < \infty$ and $0 < c < \infty$. Then as $t \to \infty$, uniformly in $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)$, β, and $\mathbf{x}^1, \ldots, \mathbf{x}^k$ such that $\alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_k \leq \beta \leq \beta_0$, $\|x_p^i - x_q^i\| \in \Gamma_t(c, \alpha_i)$ $(p \neq q)$ for each i, and $\|x_p^i - x_q^j\| \in \Gamma_t(c, \alpha_j)$ for all $i < j$, one has*

$$P\left(\#\xi_{t^\beta}^A = m\right) \to p_{n_1, \ldots, n_k; m}(\alpha_1, \ldots, \alpha_k; \beta) \quad \left(A = \left\{x_p^i \colon 1 \leq i \leq k, 1 \leq p \leq n_i\right\}\right).$$

PROOF. For $k = 1$ this is Theorem 3. We proceed by induction. Let E be the event that $\tau(x_p^i, x_q^k) \leq t^{\alpha_k}$ for some $i < k$ or for some $p \neq q$ and $i = k$. Clearly $P(E) \leq \sum P(\tau(x_p^i, x_q^j) \leq t^{\alpha_k}) \leq n_k \prod_1^k n_i \varepsilon(t; \alpha_0, \beta_0)$ by Proposition 1. So, except for a negligible term, the probability we wish to compute equals

$$\sum_l P\left(\#\xi_{t^{\alpha_k}}^B = l, \#\xi_{t^\beta}^A = m, \tilde{E}\right) \qquad \left(B = \left\{x_p^i, 1 \leq i \leq k - 1, 1 \leq p \leq n_i\right\}\right).$$

On $\tilde{E}$ the walks from $\mathbf{x}^k$ do not interact with the others. By the Markov property the last probability can be rewritten as

$$\sum_{\substack{y_1, \ldots, y_l \\ y_{l+1}, \ldots, y_{n_k + l}}} P\left(\tilde{E}, \xi_{t^{\alpha_k}}^B = \{y_1, \ldots, y_l\}, \xi_{t^{\alpha_k}}^{\mathbf{x}^k} = \{y_{l+1}, \ldots, y_{n_k + l}\}\right) P\left(\#\xi_{t^\beta - t^{\alpha_k}}^C = m\right),$$

where $C = \{y_1, \dots, y_{n_k+l}\}$ and the sum is over distinct sites y_i. By Proposition 1 all the $n_k + l$ walks are at distinct sites of order $t^{\alpha_k/2}$ at time t^{α_k} with probability $1 - \varepsilon(t; \alpha_0, \beta_0)$. Hence Proposition 3 shows that except for an error term which is at most $|\varepsilon(t; \boldsymbol{\alpha}, \beta, \mathbf{x}^1, \dots, \mathbf{x}^k)| \le \varepsilon(t; \alpha_0, \beta_0)$,

$$P\left(\#\xi^A_{t^\beta} = m\right) \approx \sum_l P\left(\#\xi^B_{t^{\alpha_k}} = l\right) p_{n_k+l, m}(\alpha_k/\beta).$$

In light of (5.4), Theorem 6 is proved by induction. □

To carry out the argument for Theorem 5 we also need a connection between the mixed moments of the Fisher–Wright diffusion and the quantities (5.3).

LEMMA 4. *Suppose* $0 \le t_1 \le t_2 \le \cdots \le t_k$ *and* $\alpha_j = e^{-t_j}$. *Then for positive integers* $n_1, n_2, \dots, n_k$ *and* $0 < \theta < 1$,

$$E_\theta\left[Y_{t_1}^{n_1} Y_{t_2}^{n_2} \cdot \cdots \cdot Y_{t_k}^{n_k}\right] = \sum_l \theta^l p_{n_k, \dots, n_1; l}(\alpha_k, \dots, \alpha_1; 1).$$

PROOF. The $k = 1$ case is (3.10). For $k \ge 2$ the Markov property gives

$$\begin{aligned} E_\theta\left[Y_{t_1}^{n_1} Y_{t_2}^{n_2} \cdot \cdots \cdot Y_{t_k}^{n_k}\right] &= E_\theta\left[E_{Y_{t_{k-1}}}\left[Y_{t_k}^{n_k}\right] Y_{t_1}^{n_1} \cdot \cdots \cdot Y_{t_{k-1}}^{n_{k-1}}\right] \\ &= \sum_j p_{n_k, j}(\alpha_k/\alpha_{k-1}) E_\theta\left[Y_{t_1}^{n_1} \cdot \cdots \cdot Y_{t_{k-2}}^{n_{k-2}} Y_{t_{k-1}}^{n_{k-1}+j}\right]. \end{aligned}$$

Proceed by induction with the aid of (5.4) to establish the claim. □

By the method of moments, and in view of Lemma 4, to finish the proof of Theorem 5 it suffices to show that for positive integers $n_1, \dots, n_k$, and $0 \le \alpha_k \le \alpha_{k-1} \le \cdots \le \alpha_1 \le 1$, as $t \to \infty$,

$$E\left[B^t(\alpha_1)^{n_1} \cdot \cdots \cdot B^t(\alpha_k)^{n_k}\right] \to \sum_l \theta^l p_{n_k, \dots, n_1; l}(\alpha_k, \dots, \alpha_1; 1). \tag{5.5}$$

Expand the left side as

$$\sum_{\substack{\mathbf{x}^1, \dots, \mathbf{x}^k: \\ \mathbf{x}^i = (x_1^i, \dots, x_{n_i}^i), \|x_p^i\| \le t^{\alpha_i/2}}} t^{-\Sigma \alpha_i n_i} P\left(\eta_t(x_j^i) \equiv 1\right) = \Sigma_1 + \Sigma_2,$$

where Σ_1 contains all $(\mathbf{x}^1, \dots, \mathbf{x}^k)$ with each $\|x_p^i - x_q^i\| \in \Gamma_t(c, \alpha_i)$ $(p \ne q)$ *and* $\|x_p^i - x_q^j\| \in \Gamma_t(c, \alpha_j)$ for $i > j$. Σ_2 contains the remaining terms. It is easy to check that the number of terms in Σ_2 is at most $C(\mathbf{n}) t^{\Sigma \alpha_i n_i}/(c^2 \log^2 t)$, and hence $\Sigma_2 = \varepsilon(t; c)$. Moreover, using (3.1) and Theorem 6,

$$\begin{aligned} \Sigma_1 &= \Sigma_1 t^{-\Sigma \alpha_i n_i} \sum_l \theta^l P\left(\#\xi_t^A = l\right) \\ &= \Sigma_1 t^{-\Sigma \alpha_i n_i} \sum_l \theta^l \left[p_{n_k, \dots, n_1; l}(\alpha_k, \dots, \alpha_1; 1) + \varepsilon(t; \mathbf{x}^1, \dots, \mathbf{x}^k)\right] \\ &= \sum_l \theta^l p_{n_k, \dots, n_1; l}(\alpha_k, \dots, \alpha_1; 1) + \varepsilon(t; \boldsymbol{\alpha}, \mathbf{n}, c), \end{aligned}$$

where $|\varepsilon(t; \alpha, n, c)| \le \varepsilon(t; \alpha)$. This gives (5.5) as desired.

Acknowledgments. We would like to thank Maury Bramson, Peter Donnelly, Phil Griffin, and Tom Kurtz, who all provided valuable clues when they were needed. Also, thanks to Craig Griffeath for lots of help with the computer simulations. Finally, we are grateful to the referee for some useful comments, especially reference [25] and other background on the stepping stone model.

REFERENCES

[1] Andjel, E. and Kipnis, C. (1985). Pointwise ergodic theorems for nonergodic interacting particle systems. Preprint.

[2] Arratia, R. (1982). Coalescing Brownian motions and the voter model on Z. Unpublished manuscript.

[3] Arratia, R. (1983). Site recurrence for annihilating random walks on Z_d. *Ann. Probab.* **11** 706–713.

[4] Bramson, M. and Griffeath, D. (1979). Renormalizing the 3-dimensional voter model. *Ann. Probab.* **7** 418–432.

[5] Bramson, M. and Griffeath, D. (1980). Clustering and dispersion rates for some interacting particle systems on $\mathbb{Z}$. *Ann. Probab.* **8** 183–213.

[6] Bramson, M. and Griffeath, D. (1980). Asymptotics for interacting particle systems on Z^d. *Z. Wahrsch. verw. Gebiete* **53** 183–196.

[7] Bramson, M., Cox, J. T., and Griffeath, D. (1985). Consolidation rates for two interacting systems in the plane. Preprint.

[8] Clifford, P. and Sudbury, A. (1973). A model for spatial conflict. *Biometrika* **60** 581–588.

[9] Cox, J. T. and Griffeath, D. (1983). Occupation time limit theorems for the voter model. *Ann. Probab.* **11** 876–893.

[10] Erdös, P. and Taylor, S. J. (1960). Some problems concerning the structure of random walk paths. *Acta Math. Acad. Sci. Hungar.* **11** 137–162.

[11] Erdös, P. and Ney, P. (1974). Some problems on random intervals and annihilating particles. *Ann. Probab.* **2** 828–839.

[12] Ethier, S. and Kurtz, T. (1981). The infinitely-many-neutral alleles diffusion model. *Adv. in Appl. Probab.* **13** 429–452.

[13] Ewens, W. (1979). *Mathematical Population Genetics. Biomathematics* **9**. Springer, New York.

[14] Griffeath, D. (1978). Annihilating and coalescing random walks on Z_d. *Z. Wahrsch. verw. Gebiete* **46** 55–65.

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

[16] Holley, R. and Liggett, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3** 643–663.

[17] Holley, R. and Stroock, D. (1979). Central limit phenomena of various interacting systems. *Ann. Math.* **110** 333–393.

[18] Jacobi, C. G. J. (1829). Fundamenta Nova Theoriae Functionum Elipticarum, Regiomontis, fratum Borntraeger. Reprinted (1881) in *Gesammelte Werke* **1**. Reimer, Berlin.

[19] Kesten, H. (1982). *Percolation Theory for Mathematicians.* Birkhäuser, Boston.

[20] Kimura, M. (1953). "Stepping stone" model of population. *Annual Rep. Nat. Inst. Genetics, Japan* **3** 62.

[21] Kimura, M. and Weiss, G. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49** 561–576.

[22] Kingman, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248.

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

[24] Malécot, G. (1948). *The Mathematics of Heredity.* English translation (1969). Freeman, San Francisco.

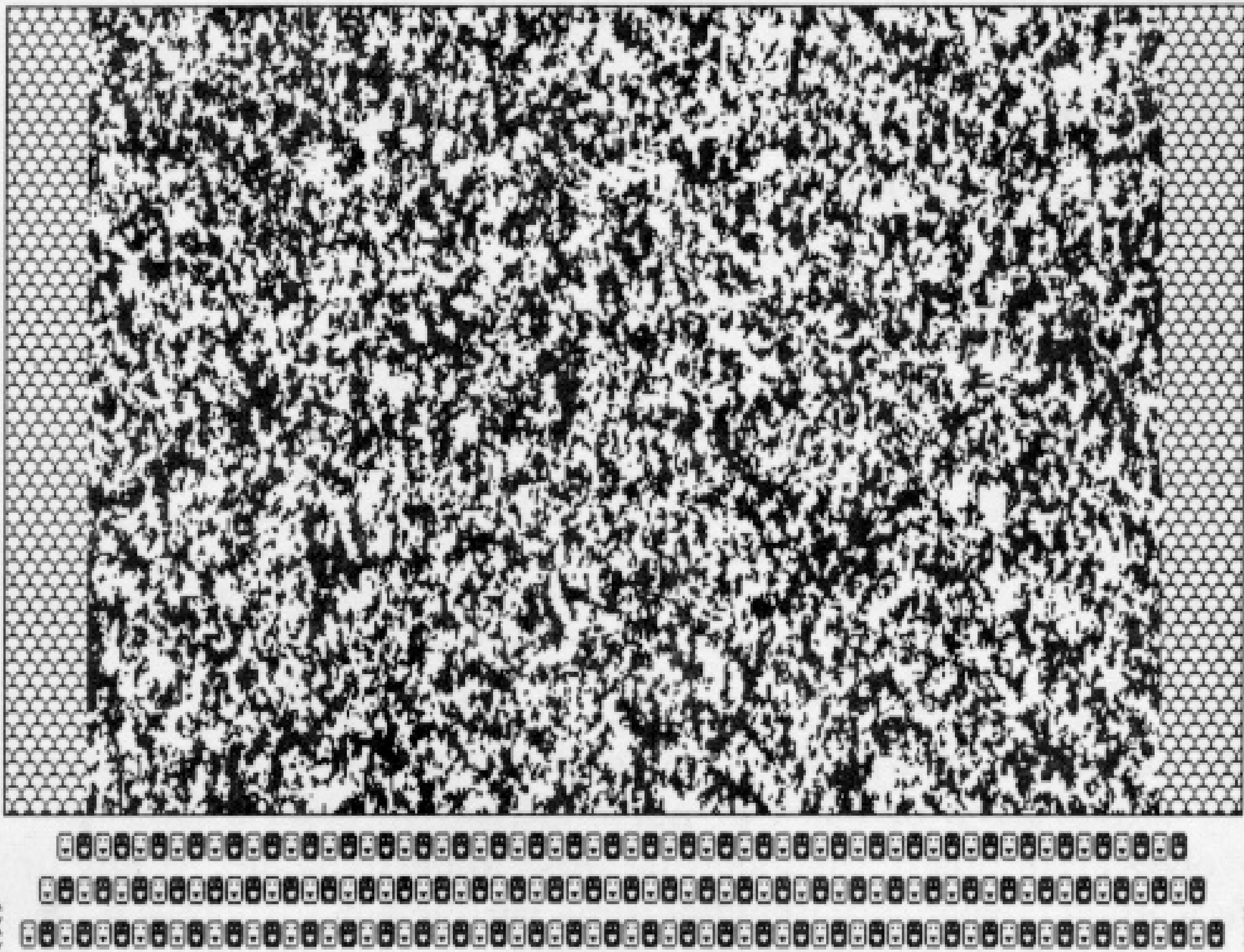
[25] Sawyer, S. (1977). Rates of consolidation in a selectively neutral migration model. *Ann. Probab.* **5** 486–493.

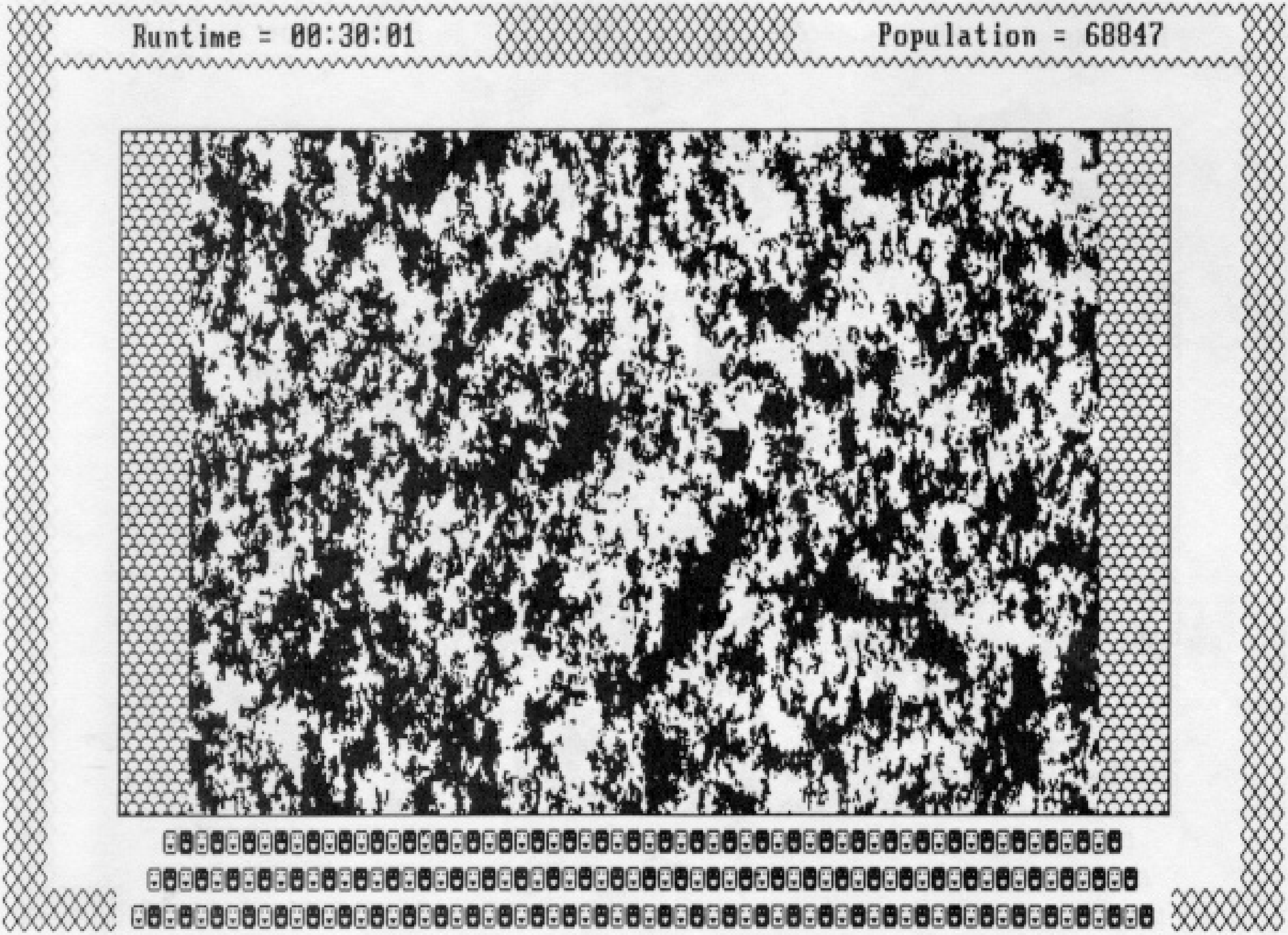
[26] SAWYER, S. (1979). A limit theorem for patch sizes in a selectively-neutral migration model. *J. Appl. Probab.* **16** 482–495.
[27] SINAI, YA. (1982). *Theory of Phase Transitions.* Pergamon, Oxford.
[28] SPITZER, F. (1976). *Principles of Random Walk.* 2nd ed. Springer, New York.
[29] TAVARÉ, 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 13210

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706

Runtime = 00:04:00 Population = 66323




Runtime = 00:30:01
Population = 68847

Runtime = 04:00:03 Population = 72311