The Annals of Probability
2001, Vol. 29, No. 3, 1001–1032

SUPER-BROWNIAN LIMITS OF VOTER MODEL CLUSTERS

BY MAURY BRAMSON,[1] J. THEODORE COX[2] AND JEAN-FRANÇOIS LE GALL

University of Minnesota, Syracuse University and Ecole Normale Supérieure

The voter model is one of the standard interacting particle systems. Two related problems for this process are to analyze its behavior, after large times t, for the sets of sites (1) sharing the same opinion as the site 0, and (2) having the opinion that was originally at 0. Results on the sizes of these sets were given by Sawyer (1979) and Bramson and Griffeath (1980). Here, we investigate the spatial structure of these sets in $d \geq 2$, which we show converge to quantities associated with super-Brownian motion, after suitable normalization. The main theorem from Cox, Durrett and Perkins (2000) serves as an important tool for these results.

1. Introduction. The voter model was introduced independently by Clifford and Sudbury in [5] (where it was called the *invasion process*) and by Holley and Liggett in [16]. It is one of the simplest *interacting particle systems* (see [15] and [20]), but one which exhibits a wide range of interesting phenomena. The process is easily described. One supposes that at each site x of the d-dimensional integer lattice $\mathbb{Z}^d$ there is a voter who randomly changes opinion. In the *two-type* model, each voter holds one of two opinions, say 0 or 1, and at rate-1 exponential random times, selects a neighbor at random according to a given jump kernel $p(x, y)$, and adopts the opinion of the neighbor at the chosen site. (Note that no change occurs if the two opinions are the same.) All voting times and neighbor selections are independent of one another. We denote the process by ξ_t, where $\xi_t(x)$ is the opinion at site x at time t, and will adopt the convention of identifying the configuration ξ_t with $\{x\colon \xi_t(x) = 1\}$, the set of sites with opinion 1. For $x \in \mathbb{Z}^d$, ξ_t^x will denote the process starting from a single 1 at the site x at time 0. The *multitype* voter model $\bar{\xi}_t$ is defined using the same dynamics as for the two-type model, but now the set of possible opinions is taken to be infinite; we will assume here that the initial opinions are all distinct. A convenient choice is to take the set of these opinions to be $\mathbb{Z}^d$, so that $\bar{\xi}_t\colon \mathbb{Z}^d \to \mathbb{Z}^d$ and $\bar{\xi}_0(x) \equiv x$.

Another basic interacting particle system is the *coalescing random walk*. Particles are assumed to execute rate-1 random walks according to some jump kernel $p(x, y)$. The movement of the particles is independent for particles at distinct sites; when particles meet, they coalesce, and afterwards move as a single particle. Unless specified otherwise, it will be assumed that there is initially a particle at each site of $\mathbb{Z}^d$. The voter model and coalescing random

Received March 2000; revised October 2000.
[1]Supported in part by NSF Grant DMS 99-71248.
[2]Supported in part by NSF Grant DMS 99-71868.
AMS 2000 *subject classifications*. Primary 60K35, 60G57; secondary 60F05, 60J80.
Key words and phrases. Voter model, super-Brownian motion, coalescing random walk.

walk are *dual processes*. In Section 2, we will give a detailed construction of both processes, using this duality to express one in terms of the other.

In this paper, we will study the limiting spatial structure of the voter model in $d \geq 2$. (The behavior for $d = 1$ is different, and will be discussed briefly at the end of the section.) These results also have analogs in terms of coalescing random walks. We first provide some background and then state the main results.

Throughout the paper, we will make certain assumptions on the jump kernel $p(x, y)$. We will assume that

$$\begin{aligned}&p(x, y) = p(0, y - x) \text{ is irreducible and symmetric, with } p(0, 0) = 0,\\ &\text{and for some } 0 < \sigma^2 < \infty, \sum_{x \in \mathbb{Z}^d} p(0, x) x^i x^j = \delta(i, j)\sigma^2\end{aligned} \tag{1.1}$$

[$\delta(i, j) = 1$ for $i = j$, and $\delta(i, j) = 0$ otherwise]. We set $\beta_2 = 2\pi\sigma^2$, and let β_d, for $d \geq 3$, be the probability that a random walk with jump kernel $p(x, y)$ starting at the origin never returns to the origin. Some of our results also require the following additional assumption:

$$\text{there exists a constant } c > 0 \text{ such that } \sum_{x \in \mathbb{Z}^d} p(0, x) e^{c|x|} < \infty. \tag{1.2}$$

Results on the sizes of the sets of interest to us were given in [22] and [4]. In [22], Sawyer studied the *patch* or *clan* of the origin π_t^0, which is the set of sites in $\bar{\xi}_t$ holding the same opinion as site 0. That is, $\pi_t^0 = \{y\colon \bar{\xi}_t(y) = \bar{\xi}_t(0)\}$. Sawyer determined the asymptotic growth of $|\pi_t^0|$, the cardinality of π_t^0. Theorem 2.1 of [22] states that, as $t \to \infty$,

$$E|\pi_t^0| \sim \begin{cases} 2\beta_2 t / \log t, & \text{in } d = 2,\\ 2\beta_d\, t, & \text{in } d \geq 3,\end{cases} \tag{1.3}$$

and

$$\frac{|\pi_t^0|}{E|\pi_t^0|} \Rightarrow \mathscr{E}(2) + \mathscr{E}'(2). \tag{1.4}$$

Here, $\mathscr{E}(2)$ and $\mathscr{E}'(2)$ are independent, exponential random variables with parameter 2, $\Rightarrow$ denotes convergence in distribution and $f(t) \sim g(t)$ means that $f(t)/g(t) \to 1$ as $t \to \infty$.

Set $p_t = P(|\xi_t^0| > 0)$. It is easy to see that $|\xi_t^0|$ is a martingale, and hence $p_t \to 0$ as $t \to \infty$. The asymptotic rate at which p_t tends to 0 was found in [4]. Theorem 1$'$ there states that, as $t \to \infty$,

$$p_t \sim \begin{cases} (\log t)/\beta_2 t, & \text{in } d = 2,\\ 1/\beta_d t, & \text{in } d \geq 3,\end{cases} \tag{1.5}$$

and

$$p_t|\hat{\xi}_t^0| \Rightarrow \mathscr{E}(1), \tag{1.6}$$

where $|\hat{\xi}_t^0|$ has the law of $|\xi_t^0|$ conditioned on the event $\{|\xi_t^0| \neq 0\}$, and $\mathscr{E}(1)$ is an exponential random variable with parameter 1. [Theorem 1$'$ in [4] was proved for the nearest neighbor random walk with $p(0, x) = 1/2d$ for $|x| = 1$; in Section 2, we will point out the minor change in reasoning needed to show that (1.5) and (1.6) hold under the more general assumption (1.1).]

It is natural to ask whether these limit theorems can be augmented with information on the *spatial structure* of ξ_t^0 and π_t^0. (This question was raised in [4].) Theorems 1 and 2 below do exactly this and express this information in terms of a measure-valued branching diffusion, *super-Brownian motion*. This process was introduced independently in [25] and [8], and has been studied extensively in recent years. (See the references in [9], [21] and [18].) We will give a brief description of it now and a more formal one in Section 3.

We start with a critical branching random walk system ζ_t. The process ζ_t models the evolution of a system of particles on $\mathbb{Z}^d$, in which each particle dies at rate r, $r > 0$, and gives birth to a new particle at the same rate. After birth, the new particle is instantly transported to a site chosen at random according to the kernel $p(x, y)$. (A particle moves only when it is born.) The number of particles at site x at time t is denoted by $\zeta_t(x)$. All death times, birth times and displacement choices are independent of one another. Super-Brownian motion is obtained by taking a *diffusion limit* of this system. This is done by speeding up time by a factor N, scaling space by $\sqrt{N}$, assigning mass $1/N$ to each particle, choosing appropriate initial conditions and letting $N \to \infty$. Here is a precise formulation.

Define a sequence of branching random walks ζ_t^N on $\mathbf{S_N} = \mathbb{Z}^d/\sqrt{N}$, with rate Nr and jump kernel $p_N(x, y) = p(x\sqrt{N}, y\sqrt{N})$, $x, y \in \mathbf{S_N}$. Assign each particle in ζ_t^N mass $1/N$, and define the corresponding *measure-valued* process X_t^N by

$$X_t^N = \frac{1}{N} \sum_{y \in \mathbf{S_N}} \zeta_t^N(y)\, \delta_y, \tag{1.7}$$

where δ_y is the unit point mass at y. Let $\mathscr{M}_f(\mathbb{R}^d)$ denote the set of finite Borel measures on $\mathbb{R}^d$, endowed with the topology of weak convergence of measures. When the (deterministic) initial measures X_0^N converge to a measure $X_0 \in \mathscr{M}_f(\mathbb{R}^d)$ as $N \to \infty$, one can show that the sequence $(X_t^N)_{t\geq 0}$ converges weakly to a continuous, measure-valued process $(X_t)_{t\geq 0}$; this limiting process is super-Brownian motion with branching rate $2r$ and diffusion coefficient σ^2. (The proof is analogous to the proof of Theorem II.5.1 of [21]). We will give a more direct definition of super-Brownian motion in Section 3.

To connect the convergence of critical branching random walks with the two-type voter model, we reformulate the voter model dynamics. Since we will be rescaling the voter model, we assume that opinions at neighboring sites are given by rate-r rather than rate-1 exponential random times. Sites with opinion 1 can be thought of as being occupied by a particle, with other sites being vacant. In this setting, a particle at x dies at rate $rV_t(x)$, where $V_t(x)$ is the *local density* of vacant sites near x,

$$V_t(x) = \sum_{y \in \mathbb{Z}^d} p(x, y) 1_{\{\xi_t(y)=0\}}.$$

Similarly, a particle at x creates a particle at rate $rV_t(x)$, with the particle being created at a vacant y at the rate $rp(y,x) = rp(x,y)$. Consequently, the voter model can be viewed as a state-dependent branching random walk in which the total branching rate of a particle at x is $2rV_t(x)$. This is the viewpoint taken in [7], where it is proved that, like the branching random walks ζ_t^N, a sequence of rescaled voter models converges to super-Brownian motion when the initial measures converge.

To be precise, let ξ_t^N denote the rate-N voter model on $\mathbf{S_N}$ with jump kernel $p_N(x,y)$, and define the mass normalizers

$$m_N = \begin{cases} N/\log N, & \text{in } d = 2, \\ N, & \text{in } d \geq 3 \end{cases} \tag{1.8}$$

and the measure-valued process X_t^N,

$$X_t^N = \frac{1}{m_N} \sum_{y \in \xi_t^N} \delta_y.$$

Theorem 1.2 of [7] states that if X_0^N converges to a measure $X_0 \in \mathscr{M}_f(\mathbb{R}^d)$ as $N \to \infty$, then

$$\left(X_t^N\right)_{t\geq 0} \Rightarrow (X_t)_{t\geq 0}, \tag{1.9}$$

where the limit process X_t is super-Brownian motion on $\mathbb{R}^d$ with branching rate $2\beta_d$ and diffusion coefficient σ^2. We note here that, for the proof of this result, it is not necessary for $N \to \infty$ over just integer values, as was assumed in [7]; for our results, we find it convenient to allow $N \to \infty$ over $\mathbb{R}_+$.

In view of (1.9), it seems plausible that, after conditioning on nonextinction of the 1 opinion of ξ_t^0 and rescaling time, space and mass, the spatial structure of ξ_t^0 should be related in some way to super-Brownian motion. This is indeed the case, and to describe this relation, we employ the family of *canonical measures* $\{R_t(x,\cdot), x \in \mathbb{R}^d, t > 0\}$ of super-Brownian motion with branching rate γ and diffusion coefficient σ^2 (see, e.g., Chapter 11 of [9]). The $R_t(x,\cdot)$ are finite measures on $\mathscr{M}_f(\mathbb{R}^d)$, which assign no mass to the zero measure, and are characterized by

$$E^\mu\left[\exp(-X_t(\phi))\right] = \exp\left(-\int_{\mathbb{R}^d}\int_{\mathscr{M}_f(\mathbb{R}^d)}\left(1 - e^{-\nu(\phi)}\right) R_t(x, d\nu)\mu(dx)\right), \tag{1.10}$$

for nonnegative measurable functions ϕ on $\mathbb{R}^d$. The notation $\nu(\phi)$ is shorthand for $\int \phi(x)\nu(dx)$; for $\mu \in \mathscr{M}_f(\mathbb{R}^d)$, X_t, under P^μ, denotes super-Brownian motion with initial state $X_0 = \mu$. We note here that $R_t(x, \mathscr{M}_f(\mathbb{R}^d)) = 2/\gamma t$. Informally, the canonical measure $R_t(x,\cdot)$ represents the law of the contribution to X_t of the descendants at time t of a single individual present at x at time 0, after normalizing the corresponding measures to compensate for "immediate" extinction. It can also be constructed as the normalized limit of the set of particles descended from a single particle in the original branching random walk system (see, e.g., Theorem II.7.2 of [21]). More precise information about canonical measures is provided in Section 3.

Our first result, Theorem 1, shows that the law of the two-type voter model ξ_t^0, conditioned on nonextinction and viewed as a measure, converges to $\widehat{R}_1(0,\cdot) = \beta_d R_1(0,\cdot)$, as $t \to \infty$, where $\{R_t(x,\cdot)\}$ is the family of canonical measures with branching rate $2\beta_d$ and diffusion coefficient σ^2. This is consistent with the exponential limit law (1.6), since the law of the total mass of a random measure distributed according to $\widehat{R}_1(0,\cdot)$ is exponential. Theorem 1 will follow as a corollary from the more general *process level* convergence result for ξ_t^0 given in Theorem 4, in Section 4, which is akin to the limit below (1.7) and to the limit (1.9). In (1.11) and elsewhere, $\mathscr{L}$ denotes law.

THEOREM 1. *Assume $d \geq 2$. As $t \to \infty$,*

$$\mathscr{L}\left(\frac{1}{m_t}\sum_{y\in\xi_t^0}\delta_{y/\sqrt{t}}\;\middle|\;\xi_t^0 \neq \varnothing\right) \Rightarrow \widehat{R}_1(0,\cdot). \tag{1.11}$$

Let d_0 denote the Hausdorff metric on nonempty compact subsets of $\mathbb{R}^d$, that is, $d_0(K,K') = d_1(K,K') + d_1(K',K)$, where

$$d_1(K,K') = \inf\{\varepsilon > 0 \colon K \subset K'_\varepsilon\},$$

and K'_ε denotes the closure of the ε-enlargement of K'. The following variant of Theorem 1 asserts that the random set $\xi_t^0/\sqrt{t}$, under $P(\cdot \mid \xi_t^0 \neq \varnothing)$, converges in distribution in the Hausdorff metric. Here, $\operatorname{supp}\mu$ denotes the closed support of the measure μ. We note that $\operatorname{supp}\mu$ is compact a.s. with respect to the measure $\widehat{R}_1(0,d\mu)$ (see Theorem IV.7 of [18]).

THEOREM 1′. *Assume $d \geq 2$, and that* (1.2) *holds. As $t \to \infty$, the law of $\xi_t^0/\sqrt{t}$ under $P(\cdot \mid \xi_t^0 \neq \varnothing)$ converges weakly to the law of* $\operatorname{supp}\mu$ *under* $\widehat{R}_1(0,d\mu)$.

Theorem 1′ will be demonstrated in Section 7. It will follow quickly from Theorem 1 once one shows that "rarefied regions," with low, nonzero densities of particles, will not occur as $t \to \infty$. Such a result is needed to ensure that the limit of $\xi_t^0/\sqrt{t}$, under $P(\cdot \mid \xi_t^0 \neq \varnothing)$, in the Hausdorff metric corresponds to that given in (1.11) (rather than the former being larger).

We next consider the patch of the origin π_t^0 for the rate-1 multitype voter model $\bar{\xi}_t$ with jump kernel $p(x,y)$. For this, we employ certain random variables $\mathscr{T}_t$, taking values in $\mathscr{M}_f(\mathbb{R}^d)$, which are characterized by

$$E[F(\mathscr{T}_t)] = \int_{\mathscr{M}_f(\mathbb{R}^d)}\int_{\mathbb{R}^d} F(\theta_z\nu)\nu(dz)R_t(0,d\nu), \qquad F \in C_b\big(\mathscr{M}_f(\mathbb{R}^d)\big). \tag{1.12}$$

[$C_b(\mathscr{M}_f(\mathbb{R}^d))$ denotes the space of continuous bounded functions on $\mathscr{M}_f(\mathbb{R}^d)$, and for $z \in \mathbb{R}^d$, θ_z denotes the shift by z; i.e., $(\theta_z\nu)(\phi) = \int \phi(y-z)\nu(dy)$.] Informally, $\mathscr{T}_t$ is the random measure obtained by viewing each measure ν from points z, which are weighted according to $\nu(dz)$ and $R_t(0,d\nu)$. (More detail on $\mathscr{T}_t$ will be given in Section 3.)

THEOREM 2. *Assume $d \geq 2$. As $t \to \infty$,*

$$\frac{1}{m_t} \sum_{y \in \pi_t^0} \delta_{y/\sqrt{t}} \Rightarrow \mathscr{T}_1. \tag{1.13}$$

As in Theorem 1′, one can rephrase Theorem 2, where the convergence in (1.13) is replaced by the convergence of the random sets $\pi_t^0/\sqrt{t}$ in the Hausdorff metric.

THEOREM 2′. *Assume $d \geq 2$ and that* (1.2) *holds. As $t \to \infty$, $\pi_t^0/\sqrt{t}$ converges in distribution to* $\operatorname{supp} \mathscr{T}_1$.

Theorem 2 follows relatively quickly from Theorem 1; it is demonstrated in Section 5. Theorem 2′ is shown in Section 7 in the same manner as Theorem 1′.

At the beginning of the section, it was mentioned that the voter model and coalescing random walk are dual processes. On account of this, one can reinterpret Theorems 1, 1′, 2 and 2′ in terms of coalescing random walks. The set ξ_t^0 for the two-type voter model is also the set of initial sites of those particles, in the coalescing random walk, which are at 0 at time t; this allows one to reinterpret Theorems 1 and 1′. Similarly, the set π_t^0 for the multitype voter model is the set of initial sites of those particles which have coalesced, by time t, with the particle starting at 0; this allows one to also reinterpret Theorems 2 and 2′. An explicit coupling of the voter model and coalescing random walk is given by their common percolation substructure, in Section 2.

In $d \geq 3$, the multitype voter model has a stationary distribution, with an infinite number of opinions, which is the limit of $\bar{\xi}_t$ as $t \to \infty$. We denote by π_∞^0 the patch of the origin for a random measure with this distribution; we view π_∞^0 as a random element of $\mathscr{M}(\mathbb{R}^d)$, the space of Radon measures μ on $\mathbb{R}^d$ (i.e., $\mu(\Gamma) < \infty$ for all compact sets Γ), endowed with the topology of vague convergence. We will later show that the random measures $\mathscr{T}_t$ converge monotonically, as $t \to \infty$, to a random measure $\mathscr{T}_\infty$ taking values in $\mathscr{M}(\mathbb{R}^d)$. The random set π_∞^0 is related to $\mathscr{T}_\infty$ in the following way.

THEOREM 3. *Assume $d \geq 3$. As $N \to \infty$,*

$$\frac{1}{N} \sum_{y \in \pi_\infty^0} \delta_{y/\sqrt{N}} \Rightarrow \mathscr{T}_\infty, \tag{1.14}$$

with respect to the topology of vague convergence on $\mathscr{M}(\mathbb{R}^d)$.

Theorem 3 is demonstrated in Section 6. It follows quickly from a variant of Theorem 2.

The results of this paper pertain to dimensions $d \geq 2$. As mentioned in the beginning of the section, the asymptotic behavior of the voter model (and coalescing random walk) is different for $d = 1$. There, the appropriate mass normalizer is $m_N = \sqrt{N}$, but the limit (1.9) does not hold without modification. One alternative is to use a sequence of jump kernels which become progressively more spread out as $N \to \infty$. This was done in Theorem 1.1 of [7]. One expects, in this case, that variants of our Theorems 1, 1′, 2 and 2′ should hold. Alternatively, one may consider the nearest neighbor voter model in $d = 1$. In this case, the endpoints of intervals of 1′s move like independent random walks, except that they "annihilate" each other when they meet. As a consequence of this fact (see [2]), the limit X_t in (1.9) exists, and is Lebesgue measure on a collection of intervals whose endpoints are annihiliating Brownian motions. It is also not difficult to argue directly that versions of our Theorems 1, 1′, 2 and 2′ hold in the nearest neighbor case. (For instance, the limit of $\xi_t^0/\sqrt{t}$, conditioned on nonextinction, is the interval between two Brownian motions, starting at the same point, conditioned not to meet before time 1.) We do not know what happens in $d = 1$ for (fixed) nonnearest neighbor kernels. However, results in [6] on the tightness of "interface" regions suggest that this case may be similar to the nearest neighbor case.

The remainder of the paper is organized as follows. Background material on the voter model and coalescing random walk is given in Section 2, and background material on super-Brownian motion is given in Section 3. Theorems 1, 1′, 2, 2′ and 3 are demonstrated in Sections 4–7, as indicated earlier. A quick application of Theorem 1′ is given in Section 8, which relates appropriate limits of the two-type voter model and coalescing random walk to a nonlinear diffusion equation.

2. The voter model and coalescing random walk.

In this section, we give the standard *graphical construction* of the voter model and its dual process, coalescing random walk. We then recall a correlation inequality from [3] and show that (1.5) and (1.6) hold under (1.1).

Let $\{\Lambda(x, y), x, y \in \mathbb{Z}^d\}$ be a family of independent Poisson point processes on $\mathbb{R}_+$, where $\Lambda(x, y)$ has intensity $p(x, y)\,ds$ (and ds denotes Lebesgue measure). The atoms of $\Lambda(x, y)$ will be the times at which the voter at x adopts the opinion of the voter at y; we indicate this by drawing an arrow from y to x at time s, for $s \in \Lambda(x, y)$. For $s < t$, we say that there is a *path up* from (y, s) to (x, t) if there is a sequence of times $s = s_0 < s_1 < s_2 \cdots < s_n \le s_{n+1} = t$ and sites $y = x_0, x_1, \dots, x_n = x$ such that:

1. For $1 \le i \le n$, there is an arrow from x_{i-1} to x_i at time s_i.
2. For $0 \le i \le n$, there are no arrows pointing towards x_i in the time interval (s_i, s_{i+1}).

There is a path down from (x, t) to (y, s) if and only if there is a path up from (y, s) to (x, t). For $t > 0$ and $x \in \mathbb{Z}^d$, define $(W_s^{x,t})_{0\le s\le t}$ by setting $W_0^{x,t} = x$ and, for $0 < s \le t$, setting $W_s^{x,t} = y$ if and only if there is a path down from (x, t) to $(y, t-s)$. It is easy to see that $W_s^{x,t}$ is a rate-1 random walk with jump kernel $p(x, y)$. Furthermore, the two walks $W_s^{x_1,t}$ and $W_s^{x_2,t}$ move independently until they meet, at which time they merge, and move together afterwards. That is, $(W_s^{x,t})_{0\le s\le t}$, $x \in \mathbb{Z}^d$, forms a coalescing random walk system.

The two-type voter model ξ_t with initial state ξ_0 is given by

$$\xi_t(x) = \xi_0\big(W_t^{x,t}\big), \tag{2.1}$$

and, in particular, ξ_t^y is the random set

$$\xi_t^y = \big\{x \colon W_t^{x,t} = y\big\}. \tag{2.2}$$

The multitype voter model $\bar{\xi}_t$ is given by the same Poisson processes via

$$\bar{\xi}_t(x) = W_t^{x,t}, \tag{2.3}$$

and π_t^x, the patch of site x at time t of the multitype voter model, is given by

$$\pi_t^x = \big\{z \colon W_t^{z,t} = W_t^{x,t}\big\}. \tag{2.4}$$

It follows easily that for any finite $A \subset \mathbb{Z}^d$ with $0 \in A$,

$$\big\{\pi_t^0 = A, W_t^{0,t} = y\big\} = \big\{\xi_t^y = A\big\}. \tag{2.5}$$

The rescaled voter models ξ_t^N and $\bar{\xi}_t^N$ may be constructed analogously using a family of independent Poisson processes $\{\Lambda^N(x, y), x, y \in \mathbf{S_N}\}$, where $\Lambda^N(x, y)$ has intensity $N p_N(x, y)\, ds$, and employing the corresponding coalescing random walks $(W_s^{N,x,t})_{0\le s\le t}$ on $\mathbf{S_N}$. Also, the analogs of (2.2) and (2.4) hold.

In the proof of Theorem 4, we will require the following correlation inequality from Lemma 1 of [3]. Recall the definition of p_t above (1.5).

LEMMA 1. *For $x \neq y$,*

$$P\big(|\xi_t^x| > 0, |\xi_t^y| > 0\big) \le P\big(|\xi_t^x| > 0\big)P\big(|\xi_t^y| > 0\big) = p_t^2. \tag{2.6}$$

We recall that the asymptotics (1.5) and (1.6) were proved, in [4], for the jump kernel $p(x, y)$ of simple symmetric random walk on $\mathbb{Z}^d$. Only Lemma 5 there makes use of this additional assumption. Display (2.7) of Lemma 2 below is the corresponding inequality, and allows us to conclude that both (1.5) and (1.6) hold under the weaker assumption (1.1).

LEMMA 2. *Let W_t denote a rate-1 random walk on $\mathbb{Z}^d$ with jump kernel $p(x, y)$ satisfying (1.1), with $W_0 = 0$. For $x \in \mathbb{Z}^d$, let $H_t(x) = P(W_s = x$ for some $s \le t)$. There exist positive constants C_d, such that for all $r \ge 2$ and $x \in \mathbb{Z}^d$ with $|x| = r$,*

$$H_{r^2}(x) \ge \begin{cases} C_2/\log r, & \text{in } d = 2,\\ C_d/r^{2-d}, & \text{in } d \ge 3.\end{cases} \tag{2.7}$$

PROOF. Let $G_t(x) = \int_0^t P(W_s = x)\, ds$. Lemma 5 of [4] relies on the inequality

$$H_t(x) \ge G_t(x)/G_t(0), \tag{2.8}$$

and on the asymptotic behavior of $G_t(x)$ for simple symmetric random walk. For $d \geq 4$, under the more general (1.1), the corresponding upper bounds on $G_t(x)$, for $x \neq 0$, require more than finite second moments on $p(x, y)$ (as noted in [26]); fortunately, the appropriate lower bounds on $G_t(x)$ do not.

We verify (2.7) under (1.1). After adaptation to continuous time, P7.9 of [23] and the remark following it imply that there exist constants $\varepsilon_d > 0$ [depending on the kernel $p(x, y)$], such that for all $r \geq 1$, $r^2/2 \leq s \leq r^2$, and $x \in \mathbb{Z}^d$ with $|x| = r$, $P(W_s = x) \geq \varepsilon_d / s^{d/2}$. Using this estimate, it follows that

$$G_{r^2}(x) \geq \varepsilon_d \int_{r^2/2}^{r^2} s^{-d/2}\, ds \geq \begin{cases} \varepsilon_2 \log 2, & \text{in } d = 2, \\ (\varepsilon_d/2) r^{2-d}, & \text{in } d \geq 3. \end{cases} \tag{2.9}$$

It also follows from P7.9 of [23], in a similar fashion, that there exist finite constants A_d such that for all $r \geq 1$, $G_{r^2}(0) \leq 1 + A_2 \log r$ for $d = 2$, and $G_{r^2}(0) \leq G_\infty(0) < \infty$ for $d \geq 3$. Substituting (2.9) and these estimates into (2.8) verifies (2.7) for $d \geq 2$, as needed. □

3. Super-Brownian motion. In this section, we summarize some of the basic properties of super-Brownian motion. For a Polish space E with Borel σ-field $\mathscr{E}$, let $\mathscr{M}(E)$ be the space of nonnegative Radon measures on $(E, \mathscr{E})$, and let $\mathscr{M}_f(E)$ [resp. $\mathscr{M}_1(E)$] be the space of finite (resp. probability) measures $\mu \in \mathscr{M}(E)$. We assign $\mathscr{M}(E)$ the topology of vague convergence, and $\mathscr{M}_f(E)$ and $\mathscr{M}_1(E)$ the topology of weak convergence. For $\mu \in \mathscr{M}_f(E)$ and functions ϕ on E, let $\mu(\phi) = \int \phi(x)\mu(dx)$ whenever the integral is well defined. Let $\mathbf{C}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ be the space of continuous functions from $\mathbb{R}_+$ to $\mathscr{M}_f(\mathbb{R}^d)$, equipped with the topology of uniform convergence on compact intervals. Let $X_t(\omega) = \omega_t$ denote the coordinate process of such a function; we will typically write X_t for $X_t(\omega)$. Also, let $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ be the Skorokhod space of cadlag functions from $\mathbb{R}_+$ to $\mathscr{M}_f(\mathbb{R}^d)$.

For $\mu \in \mathscr{M}_f(\mathbb{R}^d)$, we say that $P^\mu \in \mathscr{M}_1(\mathbf{C}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d)))$ is the law of super-Brownian motion with initial state μ, branching rate γ and diffusion coefficient σ^2 if, for all $\phi \in C_b^\infty(\mathbb{R}^d)$,

$$M_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s\left(\frac{\sigma^2 \Delta}{2} \phi\right) ds$$

is a P^μ-continuous, square-integrable martingale, with increasing process

$$\langle M(\phi) \rangle_t = \int_0^t X_s(\gamma \phi^2)\, ds.$$

See Chapter I of [21] for details on the construction of super-Brownian motion and for a proof that the above martingale problem is well posed.

Let $\mathbf{1}$ denote the function on $\mathbb{R}^d$ which is identically 1. Under P^μ, the total mass process $X_t(\mathbf{1})$ is a Feller branching diffusion process, and it is well known that

$$E^\mu\left[\exp(-\theta X_t(\mathbf{1}))\right] = \exp\left(-\frac{2\theta\mu(\mathbf{1})}{2 + \theta\gamma t}\right), \qquad \theta > 0.$$

It follows that $P^{\mu}(X_t(\mathbf{1}) > 0) = 1 - e^{-2\mu(\mathbf{1})/\gamma t}$, and hence that

$$P^{\varepsilon\delta_0}\big(X_t(\mathbf{1}) > 0\big) \sim 2\varepsilon/\gamma t \quad \text{as } \varepsilon \to 0. \tag{3.1}$$

Employing the infinite divisibility of the mass of X_t, one can show that there is a family $\{R_t(x, \cdot), x \in \mathbb{R}^d, t > 0\}$ of finite measures on $\mathscr{M}_f(\mathbb{R}^d)$ (see Chapter 11 of [9]), such that $R_t(x, \{0\}) = 0$, and for nonnegative measurable functions ϕ on $\mathbb{R}^d$,

$$E^{\mu}\big[\exp(-X_t(\phi))\big] = \exp\bigg(-\int_{\mathbb{R}^d}\int_{\mathscr{M}_f(\mathbb{R}^d)} (1 - e^{-\nu(\phi)})\, R_t(x, d\nu)\mu(dx)\bigg). \tag{3.2}$$

[This formula was given earlier as (1.10).] Equivalently, X_t under P^{μ} has the same law as $\sum X_t^i$, where $\sum \delta_{X_t^i}$ is a Poisson point process with intensity $\int R_t(x, \cdot)\,\mu(dx)$. The measures $R_t(x, \cdot)$ have total mass

$$R_t\big(x, \mathscr{M}_f(\mathbb{R}^d)\big) = 2/\gamma t, \tag{3.3}$$

and, for $\theta > 0$,

$$\int_{\mathscr{M}_f(\mathbb{R}^d)} e^{-\theta\nu(\mathbf{1})} R_t(x, d\nu) = \frac{(2/\gamma t)^2}{(2/\gamma t) + \theta}. \tag{3.4}$$

It follows from this last formula, by differentiating with respect to θ and then setting $\theta = 0$, that

$$\int_{\mathscr{M}_f(\mathbb{R}^d)} \nu(\mathbf{1})\, R_t(x, d\nu) = 1. \tag{3.5}$$

Furthermore, for any Borel set $B \subset \mathbb{R}^d$,

$$\int_{\mathscr{M}_f(\mathbb{R}^d)} \nu(B)\, R_t(x, d\nu) = \int_B n_t(x, y)\, dy, \tag{3.6}$$

where $n_t(x, y)$ is the transition density of Brownian motion in $\mathbb{R}^d$ with diffusion coefficient σ^2 (see Theorem 2.7.2 in [21]). Using (3.2), it is simple to check that

$$\lim_{\varepsilon\to 0} \varepsilon^{-1} E^{\varepsilon\delta_x}\big[1 - e^{-X_t(\phi)}\big] = \int_{\mathscr{M}_f(\mathbb{R}^d)} \big(1 - e^{-\nu(\phi)}\big)\, R_t(x, d\nu). \tag{3.7}$$

For measurable $\Upsilon \subset \mathscr{M}_f(\mathbb{R}^d)$, with $0 \notin \Upsilon$,

$$\lim_{\varepsilon\to 0} \varepsilon^{-1} P^{\varepsilon\delta_x}(X_t \in \Upsilon) = R_t(x, \Upsilon) \tag{3.8}$$

is a consequence of the Poisson representation $X_t = \sum X_t^i$ described above.

As shown in Section 4 of [13], or in Section 5 of [17] by the Brownian snake approach (see also Section II.7 of [21]), there is a σ-finite measure $\mathbf{N}_0$ on $\mathbf{C}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$, the excursion measure of super-Brownian motion with branching rate γ and diffusion coefficient σ^2, with the following properties. The measure $\mathbf{N}_0$ assigns zero mass to the zero trajectory, and to the set of trajectories with times $0 < \alpha < \beta$ such that $\omega_\alpha = 0$ and $\omega_\beta \neq 0$. For all $t > 0$ and measurable $\Upsilon \subset \mathscr{M}_f(\mathbb{R}^d)$ with $0 \notin \Upsilon$,

$$\mathbf{N}_0(X_t \in \Upsilon) = R_t(0, \Upsilon). \tag{3.9}$$

In particular, $\mathbf{N}_0(X_\alpha \neq 0) = 2/\gamma\alpha < \infty$ for any $\alpha > 0$. Thus, $\mathbf{N}_0$ restricted to $\{X_\alpha \neq 0\}$ is a finite measure. Also, for every $\delta > 0$, the process $(X_{t+\delta})_{t\geq 0}$ induced by $\mathbf{N}_0(\cdot \mid X_\delta \neq 0)$ is Markovian, with the transition kernels of super-Brownian motion having branching rate γ and diffusion coefficient σ^2. The following Poisson representation formula is useful. If $\sum_i \delta_{\omega^i}$ is a Poisson point measure on $\mathbf{C}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ with intensity $\varepsilon\mathbf{N}_0$, then

$$Y_t^{\varepsilon\delta_0} = \sum_i X_t(\omega^i), \qquad t > 0,$$

is a super-Brownian motion with initial state $\varepsilon\delta_0$.

Let $\alpha > 0$, and let F be a bounded, continuous function on $\mathbf{C}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ such that $F(\omega) = 0$ for all ω with $\omega(t) = 0$ for all $t \geq \alpha$. For such F,

$$\lim_{\varepsilon\to 0} \varepsilon^{-1} E^{\varepsilon\delta_0}\big[F\big((X_t)_{t\geq 0}\big)\big] = \mathbf{N}_0[F], \tag{3.10}$$

which is an extension of (3.8). Here, $\mathbf{N}_0[F] \overset{\text{def}}{=} \int F(\omega)\,\mathbf{N}_0(d\omega)$. (Note that for general bounded, continuous F, $\mathbf{N}_0[F]$ need not be defined.) Display (3.10) is a simple consequence of the previous representation and of properties of Poisson point measures. To see this, note that, by (3.1), (3.3) and (3.9), $P^{\varepsilon\delta_0}(X_\alpha \neq 0) \sim \varepsilon\mathbf{N}_0(X_\alpha \neq 0)$ as $\varepsilon \to 0$. Moreover, the process $(X_t)_{t\geq 0}$ is distributed under $P^{\varepsilon\delta_0}(\cdot \mid X_\alpha \neq 0)$ as the sum of two independent terms, the first term being distributed according to $\mathbf{N}_0(\cdot \mid X_\alpha \neq 0)$ and the second going to 0 in probability as $\varepsilon \to 0$. The convergence (3.10) then follows easily.

We will use a form of the Palm measures for super-Brownian motion. (See Chapter 4 of [10] for a more general theory.) The map

$$\Upsilon \mapsto \int_{\mathscr{M}_f(\mathbb{R}^d)} \int_{\mathbb{R}^d} 1_\Upsilon(\theta_z \nu)\, \nu(dz)\, R_t(0, d\nu), \tag{3.11}$$

for measurable $\Upsilon \subset \mathscr{M}_f(\mathbb{R}^d)$, defines a measure on $\mathscr{M}_f(\mathbb{R}^d)$, which by (3.5) is a probability measure. We let $\mathscr{T}_t$ denote a random measure with this law; $\mathscr{T}_t$ satisfies (1.12).

We also give an alternate, more probabilistic construction of $\mathscr{T}_t$, which will be used in the proof of Theorem 3. Let B_t^0 be a Brownian motion in $\mathbb{R}^d$ starting at 0, with diffusion coefficient σ^2. Let $\mathscr{N}(ds\, d\nu)$ be a point measure on $\mathbb{R}_+ \times \mathscr{M}_f(\mathbb{R}^d)$, such that, conditionally on the Brownian motion B^0, $\mathscr{N}$ is Poisson with intensity $\gamma ds\, R_s(B_s^0, d\nu)$, and define the random measures

$$\mathscr{I}_t = \int_0^t \int_{\mathscr{M}_f(\mathbb{R}^d)} \nu \mathscr{N}(ds\, d\nu).$$

The following result is a straightforward consequence of the Palm measure formula for superprocesses (see, e.g., page 1734 of [19]).

LEMMA 3. *For every $t > 0$, the random measures $\mathscr{T}_t$ and $\mathscr{I}_t$ have the same law.*

The equivalence of $\mathscr{T}_t$ and $\mathscr{I}_t$ is easier to see, on an intuitive level, if one considers $\mathscr{I}'_t = \int_0^t \int_{\mathscr{M}_f(\mathbb{R}^d)} (\theta_{B_t^0} \nu)\, \mathscr{N}'(ds\, d\nu)$ instead of $\mathscr{I}_t$, where $\mathscr{N}'$ is a Poisson measure with intensity $\gamma\, ds\, R_{t-s}(B_s^0, d\nu)$. In this setting, the Brownian motion B_s^0 corresponds to the historical path leading to a typical particle in the support of ν, under $R_t(0, d\nu)$. For each atom (s, ν) of the Poisson measure $\mathscr{N}'$, the measure ν corresponds to "cousins" of this particle which have common ancestry up until time s. Standard time reversal and translation arguments imply that $\mathscr{I}_t$ and $\mathscr{I}'_t$ have the same distribution.

For Theorem 3, we will also employ the random measure

$$\mathscr{I}_\infty = \int_0^\infty \int_{\mathscr{M}_f(\mathbb{R}^d)} \nu\, \mathscr{N}(ds\, d\nu).$$

Clearly,

$$\mathscr{I}_t \uparrow \mathscr{I}_\infty \qquad \text{as } t \to \infty. \tag{3.12}$$

Furthermore, for $d \geq 3$, $\mathscr{I}_\infty$ takes values in $\mathscr{M}(\mathbb{R}^d)$ (i.e., it is Radon with probability 1). If $\Gamma \subset \mathbb{R}^d$ is compact,

$$\begin{aligned} E\big[\mathscr{I}_\infty(\Gamma)\big] &= \gamma E\left[\int_0^\infty \int_{\mathscr{M}_f(\mathbb{R}^d)} \nu(\Gamma)\, R_s(B_s^0, d\nu)\, ds\right] \\ &= \gamma E\left[\int_0^\infty \int_\Gamma n_s(B_s^0, y)\, dy\, ds\right] \\ &= \gamma \int_0^\infty \int_\Gamma n_{2s}(0, y)\, dy\, ds, \end{aligned}$$

where we have used (3.6) for the second equality. For $d \geq 3$, the last integral is finite (although it is infinite for $d = 2$).

4. A process level generalization of Theorem 1. In this section, we state and prove Theorem 4, which provides the basis for the other results in the paper. Theorem 1 is, in particular, an easy consequence of Theorem 4. Recall that $\xi_t^{N,x}$ is the rate-N (two-type) voter model on $\mathbf{S_N}$ with jump kernel $p_N(x, y)$, where $\xi_t^{N,x}$ starts from a single 1, at x, at time 0. The associated random measures of $\xi_t^{N,x}$ are

$$X_t^{N,x} = \frac{1}{m_N} \sum_{y \in \xi_t^{N,x}} \delta_y,$$

where m_N is defined in (1.8).

THEOREM 4. *Assume $d \geq 2$, and let $\mathbf{N}_0$ be the excursion measure of super-Brownian motion on $\mathbb{R}^d$ with branching rate $2\beta_d$ and diffusion coefficient σ^2.*

(a) *Let* $\alpha > 0$, *and let* F *be a bounded continuous function on* $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ *such that* $F(\omega) = 0$ *for all* ω, *with* $\omega_t = 0$ *for all* $t \geq \alpha$. *Then*,

$$\lim_{N\to\infty} m_N E\Big[F\Big(\big(X_t^{N,0}\big)_{t\geq 0}\Big)\Big] = \mathbf{N}_0[F]. \tag{4.1}$$

(b) *Let* $\alpha > 0$, *and let* F *be a bounded continuous function on* $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$. *Then*,

$$\lim_{N\to\infty} E\Big[F\Big(\big(X_t^{N,0}\big)_{t\geq 0}\Big) \,\Big|\, X_\alpha^{N,0} \neq 0\Big] = \mathbf{N}_0\big[F \mid \omega_\alpha \neq 0\big]. \tag{4.2}$$

The two parts of Theorem 4 are equivalent, with part (a) containing the cleaner statement (4.1), and part (b) its more intuitive analog (4.2). The latter states that the probability measures obtained by conditioning $(X_t^{N,0})_{t\geq 0}$ on $X_\alpha^{N,0} \neq 0$ converge weakly to $\mathbf{N}_0$ conditioned on $\omega_\alpha \neq 0$. Later on in the section, we will demonstrate the theorem by showing that (b) implies (a), which is almost immediate, and then showing (b). [Part (a) also implies (b); the argument is similar to that used to prove (4.14) and (4.15) below.]

Assume that $G \in C_b(\mathscr{M}_f(\mathbb{R}^d))$ with $G(0) = 0$. For a given $\alpha > 0$, define $\widetilde{G}$ on $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ by $\widetilde{G}((X_t)_{t\geq 0}) = G(X_\alpha)$. Since $\widetilde{G}$ is a.s. continuous with respect to $\mathbf{N}_0$, it follows from (4.2) that $G(X_\alpha^{N,0})$, conditioned on $X_\alpha^{N,0} \neq 0$, converges weakly to the image of $\mathbf{N}_0[\cdot \mid \omega_\alpha \neq 0]$ under $\widetilde{G}$. By (3.3) and (3.9), this last quantity is the image of $\beta_d \alpha R_\alpha(0, \cdot)$ under G. So,

$$\lim_{N\to\infty} E\Big[G\big(X_\alpha^{N,0}\big) \mid X_\alpha^{N,0} \neq 0\Big] = \beta_d \alpha \int_{\mathscr{M}_f(\mathbb{R}^d)} G(\mu) R_\alpha(0, d\mu).$$

Theorem 1 follows from this upon substituting 1 for α and t for N. By (1.5) and (1.8) $m_N P(X_\alpha^{N,0} \neq 0) \to 1/\beta_d \alpha$ as $N \to \infty$. One can therefore also write the above limit as

$$\lim_{N\to\infty} m_N E\big[G(X_\alpha^{N,0})\big] = \int_{\mathscr{M}_f(\mathbb{R}^d)} G(\mu)\, R_\alpha(0, d\mu), \tag{4.3}$$

which is the analog (4.1). It will be applied in Section 5.

The proof of Theorem 4 is somewhat lengthy. We first summarize the basic idea and present two lemmas. For $\varepsilon > 0$, let $B_{N,\varepsilon}$ be the square in $\mathbf{S_N}$ centered at the origin of side length $b_N = (\varepsilon m_N)^{1/d}/N^{1/2}$, so that $|B_{N,\varepsilon}| \sim \varepsilon m_N$ as $N \to \infty$. Let $\eta_t^{N,\varepsilon}$ denote the voter model with initial state $B_{N,\varepsilon}$, $\eta_t^{N,\varepsilon} = \bigcup_{x\in B_{N,\varepsilon}} \xi_t^{N,x}$, and define the corresponding measures $Y_t^{N,\varepsilon}$,

$$Y_t^{N,\varepsilon} = \frac{1}{m_N} \sum_{y\in\eta_t^{N,\varepsilon}} \delta_y.$$

By the definition of $B_{N,\varepsilon}$, $Y_0^{N,\varepsilon} \to \varepsilon\delta_0$ in $\mathscr{M}_f(\mathbb{R}^d)$ as $N \to \infty$. Consequently, by (1.9),

$$\Big(Y_t^{N,\varepsilon}\Big)_{t\geq 0} \Rightarrow \Big(Y_t^{\varepsilon\delta_0}\Big)_{t\geq 0} \quad \text{as } N \to \infty, \tag{4.4}$$

where $Y_t^{\varepsilon\delta_0}$ denotes super-Brownian motion with initial state $\varepsilon\delta_0$, branching rate $2\beta_d$ and diffusion coefficient σ^2.

Roughly speaking, our strategy for proving (4.2) is to show that with high probability, when $Y_t^{N,\varepsilon} \neq 0$, there is a random site $x_N \in B_{N,\varepsilon}$ such that $Y_t^{N,\varepsilon} = X_t^{N,x_N}$. Since x_N is close to the origin, the law of X_t^{N,x_N} should be close to the law of $X_t^{N,0}$, when the latter is conditioned on nonextinction. Thus, we should be able to obtain the limiting behavior of $X_t^{N,0}$ from that of $Y_t^{N,\varepsilon}$, by letting $\varepsilon \to 0$ and $N \to \infty$.

Let $S_t^{N,\varepsilon}$ be the set of surviving family lines at time t from $\eta_t^{N,\varepsilon}$,

$$S_t^{N,\varepsilon} = \Big\{x \in B_{N,\varepsilon}\colon |\xi_t^{N,x}| > 0\Big\}. \tag{4.5}$$

Our first lemma shows that one may neglect the possibility that there are two or more surviving family lines at a fixed rescaled time.

LEMMA 4. *For any* $\delta > 0$,

$$P\Big(|S_\delta^{N,\varepsilon}| \geq 2\Big) \leq |B_{N,\varepsilon}|^2 p_{\delta N}^2 \sim \big(\varepsilon/\delta\beta_d\big)^2 \tag{4.6}$$

as $N \to \infty$.

PROOF. By a simple decomposition and Lemma 1,

$$\begin{aligned} P(|S_\delta^{N,\varepsilon}| \geq 2) &= P\left(\bigcup_{\substack{x,y\in B_{N,\varepsilon}\\ x\neq y}} \Big\{|\xi_\delta^{N,x}| > 0, |\xi_\delta^{N,y}| > 0\Big\}\right) \\ &\leq \sum_{\substack{x,y\in B_{N,\varepsilon}\\ x\neq y}} P\Big(|\xi_\delta^{N,x}| > 0, |\xi_\delta^{N,y}| > 0\Big) \\ &\leq |B_{N,\varepsilon}|^2 p_{\delta N}^2. \end{aligned}$$

Now apply (1.5). □

We will need certain bounds on the total mass process of super-Brownian motion. The total mass process $X_t(\mathbf{1})$ is a Feller diffusion U_t, defined by

$$dU_t = \sqrt{\gamma U_t}\, dB_t, \tag{4.7}$$

where B_t is a standard Brownian motion on $\mathbb{R}$. Let U_t^ε denote this diffusion with initial value $\varepsilon > 0$.

LEMMA 5. *For* $\delta > 0$ *and* $\alpha > 0$, *let*

$$c_\delta(\alpha) = \limsup_{\varepsilon\to 0} \varepsilon^{-1} E\bigg[\bigg(\sup_{0\leq t\leq\delta} U_t^\varepsilon\bigg) \wedge U_\alpha^\varepsilon \wedge 1\bigg]. \tag{4.8}$$

Then, $c_\delta(\alpha) \to 0$ *as* $\delta \to 0$.

PROOF. The argument is based on the following basic properties of U_t^ε: (1) U_t^ε is a Markov process and (2) U_t^ε is a square integrable continuous martingale. We also use the following formulas that can be derived from the Laplace transform of U_δ^ε, which is given above (3.1). For $\delta > 0$,

$$E\big[(U_\delta^\varepsilon)^2\big] = \varepsilon^2 + \gamma\delta\varepsilon \tag{4.9}$$

and

$$P\big(U_\delta^\varepsilon > 0\big) = 1 - e^{-2\varepsilon/\gamma\delta}. \tag{4.10}$$

By the Markov property at time δ and (4.10), we have, for $\delta < \alpha/2$,

$$\begin{aligned}
E\bigg[\Big(\sup_{0\le t\le\delta} U_t^\varepsilon\Big) \wedge U_\alpha^\varepsilon \wedge 1\bigg] &\le E\bigg[\Big(\sup_{0\le t\le\delta} U_t^\varepsilon\Big) 1_{\{U_\alpha^\varepsilon > 0\}}\bigg] \\
&= E\bigg[\Big(\sup_{0\le t\le\delta} U_t^\varepsilon\Big) P\Big(U_\alpha^\varepsilon > 0 \mid U_\delta^\varepsilon\Big)\bigg] \\
&= E\bigg[\Big(\sup_{0\le t\le\delta} U_t^\varepsilon\Big)\Big(1 - e^{-2U_\delta^\varepsilon/\gamma(\alpha-\delta)}\Big)\bigg] \\
&\le \frac{4}{\gamma\alpha} E\bigg[\Big(\sup_{0\le t\le\delta} U_t^\varepsilon\Big)^2\bigg].
\end{aligned}$$

By Doob's inequality, this is

$$\le \frac{16}{\gamma\alpha} E\Big[(U_\delta^\varepsilon)^2\Big].$$

The lemma follows from this bound and (4.9). □

Before starting the proof of Theorem 4, we make a few observations concerning weak convergence on $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$. Recall that $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$, with the Skorokhod metric, is a complete metric space. Note, for this, that the topology of weak convergence on the space $\mathscr{M}_f(\mathbb{R}^d)$ is given by the metric d,

$$d(\mu, \nu) = \sup_{f \in B_L(\mathbb{R}^d)} |\mu(f) - \nu(f)|, \tag{4.11}$$

where $B_L(\mathbb{R}^d)$ denotes the set of all nonnegative functions on $\mathbb{R}^d$ which are bounded by 1, and are Lipschitz with Lipschitz constant at most 1. (See Problems 3.11.2 and 9.5.6 in [14].) Let $\mathscr{F}$ denote the set of Lipschitz functions $F(\omega)$ on $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$, with $0 \le F(\omega) \le 1$, which depend only on $(\omega_t, 0 \le t \le K_F)$ for some $K_F > 0$. By Theorem 3.4.5 of [14], $\mathscr{F}$ is convergence determining on $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ [i.e., for probability measures Q and Q_N, $\int F\,dQ_N \to \int F\,dQ$ as $N \to \infty$, for all $F \in \mathscr{F}$, implies that $Q_N \Rightarrow Q$], since $\mathscr{F}$ strongly separates points. We also note that the Skorokhod metric on $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$, when restricted to functions that only differ on $[0, K]$, $K > 0$, is bounded above by the corresponding uniform metric on $[0, K]$. It follows that, for each $F \in \mathscr{F}$, there exists a constant $C_F \geq 1$, such that, for every $\omega, \omega' \in D(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$,

$$|F(\omega) - F(\omega')| \leq C_F \sup_{0 \leq t \leq K_F} d\big(\omega(t), \omega'(t)\big). \tag{4.12}$$

We will employ measurable functions F satisfying (4.12) and $0 \leq F(\omega) \leq 1$, with the further restriction given in (4.13), in the proof of Theorem 4. We will also employ related sets of convergence determining functions on $\mathscr{M}_f(\mathbb{R}^d)$ in Theorems 2 and 3 in Sections 5 and 6.

PROOF OF THEOREM 4. It is easy to see that (4.1) follows from (4.2). We first note that $P(X_\alpha^{N,0} \neq 0) = P(|\xi_\alpha^{N,0}| \neq 0) \sim 1/\alpha\beta_d m_N$ as $N \to \infty$, by (1.5) and (1.8). On the other hand, by (3.3) and (3.9), $\mathbf{N}_0(\omega_\alpha \neq 0) = R_\alpha(0, \mathscr{M}_f(\mathbb{R}^d)) = 1/\alpha\beta_d$. So, for F satisfying the assumptions of part (a), (4.2) implies (4.1).

The remainder of the proof is devoted to demonstrating (4.2) for any measurable function F on $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ satisfying $0 \leq F \leq 1$ and condition (4.12). It suffices to further restrict F so that

$$F(\omega) \leq C_F\, \omega_\alpha(\mathbf{1}), \tag{4.13}$$

where $\alpha > 0$ is as in (4.2). Note that (4.13) implies the condition on F given in part (a) of Theorem 4, that $F(\omega) = 0$ for all ω, with $\omega_t = 0$ for all $t \geq \alpha$. To see that the additional restriction (4.13) is justified, we argue as follows.

Suppose that (4.2) holds under (4.13). Let $F_n(\omega) = F(\omega)g_n(\omega)$, where $g_n(\omega) = 1 \wedge (n\omega_\alpha(\mathbf{1}))$. Since, for each $n > 0$, $F_n(\omega) \leq n\omega_\alpha(\mathbf{1})$, F_n satisfies (4.13), and so

$$\lim_{N\to\infty} E\Big[F_n\big((X_t^{N,0})_{t\geq 0}\big) \mid X_\alpha^{N,0} \neq 0\Big] = \mathbf{N}_0[F_n \mid \omega_\alpha \neq 0].$$

Since $F_n \leq F$, and $F_n \to F 1_{\{\omega_\alpha \neq 0\}}$ as $n \to \infty$, monotone convergence implies

$$\liminf_{N\to\infty} E\Big[F\big((X_t^{N,0})_{t\geq 0}\big) \mid X_\alpha^{N,0} \neq 0\Big] \geq \mathbf{N}_0\big[F \mid \omega_\alpha \neq 0\big]. \tag{4.14}$$

Replacing F with $1 - F$ in (4.14), we obtain

$$\liminf_{N\to\infty} E\Big[1 - F\big((X_t^{N,0})_{t\geq 0}\big) \mid X_\alpha^{N,0} \neq 0\Big] \geq \mathbf{N}_0\big[1 - F \mid \omega_\alpha \neq 0\big]$$

and hence,

$$\limsup_{N\to\infty} E\Big[F\Big(\big(X_t^{N,0}\big)_{t\geq 0}\Big) \mid X_\alpha^{N,0} \neq 0\Big] \leq \mathbf{N}_0\big[F \mid \omega_\alpha \neq 0\big]. \tag{4.15}$$

Together, (4.14) and (4.15) imply (4.2).

In the remainder of the proof, it will be more convenient to employ the format of (4.1), instead of (4.2), but with the restrictions on F given above. That is, we will prove that, for functions F on $\mathbf{D}(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}^d))$ satisfying $0 \leq F \leq 1$ and conditions (4.12) and (4.13),

$$\lim_{N\to\infty} m_N E\Big[F\Big(\big(X_t^{N,0}\big)_{t\geq 0}\Big)\Big] = \mathbf{N}_0[F]. \tag{4.16}$$

Given (4.16), (4.2), for this class of functions F, follows easily by again using the estimates on $P(X_\alpha^{N,0} \neq 0)$ and $\mathbf{N}_0(\omega_\alpha \neq 0)$ in the first paragraph of the proof.

In order to demonstrate (4.16), we will employ the following six displays, (4.17)–(4.22). For these displays, recall that $Y_t^{\varepsilon\delta_0}$ denotes super-Brownian motion with branching rate $2\beta_d$ and diffusion coefficient σ^2, $Y_t^{N,\varepsilon}$ is the normalized voter model process defined above (4.4) and U_t^ε is a Feller branching diffusion started at ε. The function F is assumed to satisfy the conditions specified in the previous paragraph, and we set $F_x = F\circ\theta_x$, where $\theta_x\omega = (\theta_x\omega_t)_{t\geq 0}$; $\varepsilon > 0$ and $\delta \in (0,\alpha)$ are also assumed. We will first demonstrate (4.16), assuming (4.17)–(4.22), and will afterwards justify these displays. They are

$$\lim_{N\to\infty} E\Big[F\Big((Y_t^{N,\varepsilon})_{t\geq 0}\Big)\Big] = E\Big[F\Big((Y_t^{\varepsilon\delta_0})_{t\geq 0}\Big)\Big], \tag{4.17}$$

$$\limsup_{N\to\infty}\left|E\Big[F\Big((Y_t^{N,\varepsilon})_{t\geq 0}\Big)\Big] - E\Big[F\Big((Y_t^{N,\varepsilon})_{t\geq 0}\Big)1_{\{|S_\delta^{N,\varepsilon}|=1\}}\Big]\right| \leq \left(\frac{\varepsilon}{\delta\beta_d}\right)^2, \tag{4.18}$$

$$\begin{aligned}\lim_{N\to\infty} E\Bigg[\Bigg|&F\Big((Y_t^{N,\varepsilon})_{t\geq 0}\Big)1_{\{|S_\delta^{N,\varepsilon}|=1\}}\\ &- \sum_{x\in B_{N,\varepsilon}} F_x\Big((Y_t^{N,\varepsilon})_{t\geq 0}\Big)\,1_{\{S_\delta^{N,\varepsilon}=\{x\}\}}\Bigg|\Bigg] = 0,\end{aligned} \tag{4.19}$$

$$\begin{aligned}\limsup_{N\to\infty}&\left|E\left[\sum_{x\in B_{N,\varepsilon}}\Big[F_x\Big((Y_t^{N,\varepsilon})_{t\geq 0}\Big) - F_x\Big((X_t^{N,x})_{t\geq 0}\Big)\Big]1_{\{S_\delta^{N,\varepsilon}=\{x\}\}}\right]\right|\\ &\leq C_F E\left[\left(\sup_{0\leq t\leq\delta} U_t^\varepsilon\right)\wedge U_\alpha^\varepsilon\wedge 1\right],\end{aligned} \tag{4.20}$$

$$\begin{aligned}\limsup_{N\to\infty}\Bigg|E\Bigg[&\sum_{x\in B_{N,\varepsilon}} F_x\Big((X_t^{N,x})_{t\geq 0}\Big)\,1_{\{S_\delta^{N,\varepsilon}=\{x\}\}}\\ &-|B_{N,\varepsilon}|F\Big((X_t^{N,0})_{t\geq 0}\Big)\Bigg]\Bigg| \leq \left(\frac{\varepsilon}{\delta\beta_d}\right)^2,\end{aligned} \tag{4.21}$$

$$\mathbf{N}_0[F] = \lim\nolimits_{\varepsilon\to 0}\varepsilon^{-1}E\Big[F\big((Y_t^{\varepsilon\delta_0})_{t\geq 0}\big)\Big]. \tag{4.22}$$

Combining (4.17)–(4.21), one obtains

$$\begin{aligned}\limsup_{N\to\infty}&\left||B_{N,\varepsilon}|\,E\Big[F\Big((X_t^{N,0})_{t\geq 0}\Big)\Big] - E\Big[F\Big((Y_t^{\varepsilon\delta_0})_{t\geq 0}\Big)\Big]\right|\\ &\leq\, 2\left(\frac{\varepsilon}{\delta\beta_d}\right)^2 + C_F E\left[\left(\sup_{0\leq t\leq\delta} U_t^\varepsilon\right)\wedge U_\alpha^\varepsilon\wedge 1\right].\end{aligned}$$

Consequently,

$$\begin{aligned}\limsup_{N\to\infty}&\left|\varepsilon^{-1}|B_{N,\varepsilon}|\,E\Big[F\Big(\big(X_t^{N,0}\big)_{t\ge0}\Big)\Big]-\mathbf{N}_0[F]\right|\\&\le\frac{2\varepsilon}{(\delta\beta_d)^2}+\left|\varepsilon^{-1}E\Big[F\Big(\big(Y_t^{\varepsilon\delta_0}\big)_{t\ge0}\Big)\Big]-\mathbf{N}_0[F]\right|\\&\quad+C_F\varepsilon^{-1}E\bigg[\bigg(\sup_{0\le t\le\delta}U_t^\varepsilon\bigg)\wedge U_\alpha^\varepsilon\wedge 1\bigg].\end{aligned}$$

Since $m_N\sim\varepsilon^{-1}|B_{N,\varepsilon}|$, we can replace $\varepsilon^{-1}|B_{N,\varepsilon}|$ with m_N on the left side of the above inequality, which then becomes independent of ε. Letting ε go to 0 on the right side and defining $c_\delta(\alpha)$ as in Lemma 5, (4.22) implies that

$$\limsup_{N\to\infty}\left|m_NE\Big[F\Big(\big(X_t^{N,0}\big)_{t\ge0}\Big)-\mathbf{N}_0[F]\right|\le C_Fc_\delta(\alpha).$$

By Lemma 5, the right side goes to 0 as $\delta\to0$. We have thus proved (4.16), assuming (4.17)–(4.22), for our restricted class of functions F; as explained earlier, this implies (4.2) for general F.

We need to justify (4.17)–(4.22). The limit (4.22) is (3.10) and (4.17) follows from (4.4), since F which satisfy (4.12) are continuous on $\mathbf{C}(\mathbb{R}_+,\mathscr{M}_f(\mathbb{R}^d))$. We next show (4.18). Since, by (4.13), $F(0)=0$,

$$E\Big[F\Big(\big(Y_t^{N,\varepsilon}\big)_{t\ge0}\Big)\Big]=E\Big[F\Big(\big(Y_t^{N,\varepsilon}\big)_{t\ge0}\Big)1_{\{|S_\delta^{N,\varepsilon}|\ge1\}}\Big].$$

The inequality (4.18) follows from this and Lemma 4.

In order to show (4.19), we note that by (4.12),

$$|F(\omega)-F_x(\omega)|\le C_F\sup_{0\le t\le K}d(\omega_t,\theta_x\omega_t)\le C_F\,|x|\sup_{0\le t\le K}\omega_t(\mathbf{1}).$$

These inequalities, the bound $F\le1$ and the fact that the events $\{S_\delta^{N,\varepsilon}=\{x\}\}$, $x\in B_{N,\varepsilon}$ are disjoint imply that

$$\begin{aligned}E&\bigg[\bigg|\sum_{x\in B_{N,\varepsilon}}\Big[F\Big(\big(Y_t^{N,\varepsilon}\big)_{t\ge0}\Big)-F_x\Big(\big(Y_t^{N,\varepsilon}\big)_{t\ge0}\Big)\Big]1_{\{S_\delta^{N,\varepsilon}=\{x\}\}}\bigg|\bigg]\\&\le C_FE\bigg[\bigg(db_N\sup_{0\le t\le K}Y_t^{N,\varepsilon}(\mathbf{1})\bigg)\wedge1\bigg].\end{aligned}$$

(Recall that $B_{N,\varepsilon}$ has side length b_N; here d is its dimension.) Since $b_N\to0$, it follows from (4.4) and bounded convergence that the right side goes to 0 as $N\to\infty$. The limit (4.19) follows from this and the decomposition

$$E\Big[F\Big(\big(Y_t^{N,\varepsilon}\big)_{t\ge0}\Big)1_{\{|S_\delta^{N,\varepsilon}|=1\}}\Big]=\sum_{x\in B_{N,\varepsilon}}E\Big[F\Big(\big(Y_t^{N,\varepsilon}\big)_{t\ge0}\Big)1_{\{S_\delta^{N,\varepsilon}=\{x\}\}}\Big].$$

For (4.20), note that, on the event $\{S_\delta^{N,\varepsilon} = \{x\}\}$, $X_t^{N,x} = Y_t^{N,\varepsilon}$ holds for all $t \geq \delta$. Also, $X_t^{N,x} \leq Y_t^{N,\varepsilon}$ always holds for all t. From the assumptions (4.12) and (4.13) on F and $0 \leq F \leq 1$, it follows that, for every $x \in B_{N,\varepsilon}$,

$$\begin{aligned}\left|F_x\Big(\big(Y_t^{N,\varepsilon}\big)_{t\geq 0}\Big) - F_x\Big(\big(X_t^{N,x}\big)_{t\geq 0}\Big)\right| 1_{\{S_\delta^{N,\varepsilon}=\{x\}\}} \\ \leq C_F\left(\sup_{0\leq t\leq\delta}\Big(Y_t^{N,\varepsilon}(\mathbf{1})\Big) \wedge Y_\alpha^{N,\varepsilon}(\mathbf{1}) \wedge 1\right) 1_{\{S_\delta^{N,\varepsilon}=\{x\}\}}.\end{aligned}$$

Since the events $\{S_\delta^{N,\varepsilon} = \{x\}\}$ are disjoint, it follows from this, that

$$\begin{aligned}\limsup_{N\to\infty}\left|E\left[\sum_{x\in B_{N,\varepsilon}}\left[F_x\Big(\big(Y_t^{N,\varepsilon}\big)_{t\geq 0}\Big) - F_x\Big(\big(X_t^{N,x}\big)_{t\geq 0}\Big)\right]1_{\{S_\delta^{N,\varepsilon}=\{x\}\}}\right]\right| \\ \leq C_F \limsup_{N\to\infty} E\left[\sup_{0\leq t\leq\delta}\Big(Y_t^{N,\varepsilon}(\mathbf{1})\Big) \wedge Y_\alpha^{N,\varepsilon}(\mathbf{1}) \wedge 1\right].\end{aligned}$$

Together with (4.4), this implies (4.20).

We still need to show (4.21). The reasoning is almost the same as that for (4.18). Since $F(\omega) = 0$ if $\omega_\alpha = 0$,

$$E\left[F\Big(\big(X_t^{N,x}\big)_{t\geq 0}\Big)\right] = E\left[F\Big(\big(X_t^{N,x}\big)_{t\geq 0}\Big)1_{\{|S_\delta^{N,\varepsilon}|\geq 1\}}\right],$$

for each x. The same simple decomposition as in Lemma 4 therefore shows that

$$\begin{aligned}\left|E\left[\sum_{x\in B_{N,\varepsilon}} F_x\Big(\big(X_t^{N,x}\big)_{t\geq 0}\Big)1_{\{S_\delta^{N,\varepsilon}=\{x\}\}} - \sum_{x\in B_{N,\varepsilon}} F_x\Big(\big(X_t^{N,x}\big)_{t\geq 0}\Big)\right]\right| \\ \leq |B_{N,\varepsilon}|^2 p_{\delta N}^2;\end{aligned}$$

the right side $\sim (\varepsilon/\delta\beta_d)^2$ for large N. The limit (4.21) follows from this and

$$E\left[F_x\Big(\big(X_t^{N,x}\big)_{t\geq 0}\Big)\right] = E\left[F\Big(\big(X_t^{N,0}\big)_{t\geq 0}\Big)\right]. \qquad \square$$

5. Convergence of the patch of the origin. We introduce the notation

$$\pi_t^{N,0} = \left\{y\colon W_t^{N,y,t} = W_t^{N,0,t}\right\}, \qquad \Pi_t^{N,0} = \frac{1}{m_N}\sum_{y\in\pi_t^{N,0}}\delta_y,$$

where $(W_s^{N,y,t})_{0\leq s\leq t}$ are the coalescing random walks with jump rates N on $\mathbf{S_N}$, which were introduced in Section 2. Thus, $\pi_t^{N,0}$ is the patch of the origin after scaling time by N and space by $\sqrt{N}$, and $\Pi_t^{N,0}$ is the corresponding measure after normalization by m_N. In this section, we prove that for all $t > 0$ and $F \in C_b(\mathscr{M}_f(\mathbb{R}^d))$,

$$\lim_{N\to\infty} E\Big[F\big(\Pi_t^{N,0}\big)\Big] = E[F(\mathscr{T}_t)], \tag{5.1}$$

where $\mathscr{T}_t$ is given by (3.11) [see also (1.12)]. Theorem 2 follows by substituting 1 for t and t for N in (5.1). The proof of (5.1) uses (4.3).

The first step is to derive the following representation.

LEMMA 6.

$$E\Big[F\big(\Pi_t^{N,0}\big)\Big] = m_N E\bigg[\int_{\mathbb{R}^d} F\Big(\theta_z X_t^{N,0}\Big)\, X_t^{N,0}(dz)\bigg], \qquad t \geq 0. \tag{5.2}$$

PROOF. Let $\mathscr{A}_N$ be the collection of finite subsets of $\mathbf{S_N}$. As in (2.5), for all $y \in \mathbf{S_N}$ and $A \in \mathscr{A}_N$, with $0 \in A$, the events $\{\pi_t^{N,0} = A,\ W_t^{N,0,t} = -y\}$ and $\{\xi_t^{N,-y} = A\}$ coincide ($-y$ is more convenient than y for the next calculation). Using this and $P(\xi_t^{N,-y} = A) = P(\xi_t^{N,0} = A + y)$, we have

$$\begin{aligned}
E\Big[F\Big(\Pi_t^{N,0}\Big)\Big] &= \sum_{A\in\mathscr{A}_N}\sum_{y\in\mathbf{S_N}} 1_A(0) E\Big[F\Big(\Pi_t^{N,0}\Big) 1_{\{\pi_t^{N,0}=A,\, W_t^{N,0,t}=-y\}}\Big]\\
&= \sum_{A\in\mathscr{A}_N}\sum_{y\in\mathbf{S_N}} 1_A(0) F\bigg(m_N^{-1}\sum_{x\in A}\delta_x\bigg) P\Big(\xi_t^{N,-y} = A\Big)\\
&= \sum_{A\in\mathscr{A}_N}\sum_{y\in\mathbf{S_N}} 1_A(0) F\bigg(m_N^{-1}\sum_{x\in A}\delta_x\bigg) P\big(\xi_t^{N,0} = A + y\big).
\end{aligned}$$

Changing variables, we obtain

$$E\Big[F\Big(\Pi_t^{N,0}\Big)\Big] = \sum_{A\in\mathscr{A}_N}\sum_{y\in\mathbf{S_N}} 1_A(y) F\bigg(\theta_y\bigg(m_N^{-1}\sum_{x\in A}\delta_x\bigg)\bigg) P\Big(\xi_t^{N,0} = A\Big)$$

(recall that $\theta_y\mu$ is the shift of μ by y). Consequently,

$$E\Big[F\big(\Pi_t^{N,0}\big)\Big] = E\bigg[\sum_{y\in\mathbf{S_N}} F\big(\theta_y X_t^{N,0}\big)\, \xi_t^{N,0}(y)\bigg]$$

and since $\xi_t^{N,0}(y) = m_N X_t^{N,0}(\{y\})$, (5.2) follows. $\square$

Letting $G(\mu) = \int_{\mathbb{R}^d} F(\theta_z\mu)\,\mu(dz)$, we can rewrite (5.2) in the form

$$E\Big[F\Big(\Pi_t^{N,0}\Big)\Big] = m_N E\Big[G\Big(X_t^{N,0}\Big)\Big]. \tag{5.3}$$

Employing this and (1.12), it suffices to show that

$$\lim_{N\to\infty} m_N E\Big[G\Big(X_t^{N,0}\Big)\Big] = \int_{\mathscr{M}_f(\mathbb{R}^d)} G(\mu) R_t(0, d\mu) \tag{5.4}$$

in order to show (5.1).

In (5.1), and hence in (5.4), it suffices to also assume (by reasoning analogous to that in the paragraph before the proof of Theorem 4) that $0 \le F \le 1$, and F is Lipschitz with Lipschitz constant at most 1. We claim that, under these conditions, G is continuous. To see this, note that

$$\big|F(\theta_z\mu) - F(\theta_z\nu)\big| \le d(\theta_z\mu, \theta_z\nu) = d(\mu,\nu) \tag{5.5}$$

and

$$|F(\theta_z\nu) - F(\theta_{z'}\nu)| \le |z - z'|\nu(\mathbf{1}). \tag{5.6}$$

Applying (5.5) to the first integral below, and (5.6) together with (4.11) [for $f(z) = F(\theta_z\nu)/(1 \vee \nu(\mathbf{1})) \in B_L(\mathbb{R}^d)$] to the second integral, one obtains that

$$\begin{aligned} |G(\mu) - G(\nu)| &\le \int |F(\theta_z\mu) - F(\theta_z\nu)|\,\mu(dz) \\ &\qquad + \left|\int F(\theta_z\nu)\,\mu(dz) - \int F(\theta_z\nu)\,\nu(dz)\right| \\ &\le \Big[\mu(\mathbf{1}) + \big(1 \vee \nu(\mathbf{1})\big)\Big] d(\mu,\nu). \end{aligned}$$

Thus, G is continuous.

In order to demonstrate (5.4), we would like to apply (4.3) to G. It is easy to see that $G(0) = 0$. It is not bounded, however, and so we set $G_n(\mu) = n \wedge G(\mu)$. Applying (4.3) to G_n, one obtains

$$\lim_{N\to\infty} m_N E\Big[G_n\big(X_t^{N,0}\big)\Big] = \int_{\mathscr{M}_f(\mathbb{R}^d)} G_n(\mu)\,R_t(0,d\mu). \tag{5.7}$$

By monotone convergence, the right side above converges to $\int_{\mathscr{M}_f(\mathbb{R}^d)} G(\mu) \times R_t(0,d\mu)$ as $n \to \infty$. Since $G_n \le G$, this implies that

$$\liminf_{N\to\infty} m_N E\Big[G\big(X_t^{N,0}\big)\Big] \ge \int_{\mathscr{M}_f(\mathbb{R}^d)} G(\mu)\,R_t(0,d\mu). \tag{5.8}$$

If F is replaced with $1 - F$, then $G(\mu)$ is replaced with $\widehat{G}(\mu) = \mu(\mathbf{1}) - G(\mu)$ in (5.8). Note that $m_N E[X_t^{N,0}(\mathbf{1})]$ and $\int \mu(\mathbf{1})\,R_t(0,d\mu)$ both equal 1, by (3.5). Consequently,

$$\limsup_{N\to\infty} m_N E\Big[G\big(X_t^{N,0}\big)\Big] \le \int_{\mathscr{M}_f(\mathbb{R}^d)} G(\mu)\,R_t(0,d\mu).$$

Together with (5.8), this implies (5.4).

6. Proof of Theorem 3. In this section we assume that $d \ge 3$, and prove Theorem 3. It will be convenient to introduce another family of rate-N coalescing random walks on $\mathbf{S_N}$, $\{(W_s^{N,x})_{s\ge 0}, x \in \mathbf{S_N}\}$, where, for each $t > 0$, the law of $\{(W_s^{N,x})_{0\le s\le t}, x \in \mathbf{S_N}\}$ is the same as that of $\{(W_s^{N,x,t})_{0\le s\le t}, x \in \mathbf{S_N}\}$. (This extension allows pathwise comparisons between $W_s^{N,x}$ at different s.) Let $\tau^N(x) = \inf\{t\colon W_t^{N,x} = W_t^{N,0}\}$ [where $\tau^N(x) = \infty$ if the set is void] and define

$$\bar{\pi}_t^{N,0} = \big\{y\colon \tau^N(y) \le t\big\}, \qquad \bar{\pi}_\infty^{N,0} = \big\{y\colon \tau^N(y) < \infty\big\} \tag{6.1}$$

and the associated measures

$$\overline{\Pi}_t^{N,0} = \frac{1}{N} \sum_{y \in \bar{\pi}_t^{N,0}} \delta_y, \qquad \overline{\Pi}_\infty^{N,0} = \frac{1}{N} \sum_{y \in \bar{\pi}_\infty^{N,0}} \delta_y. \tag{6.2}$$

Note that

$$\overline{\Pi}_t^{N,0} \stackrel{\text{(d)}}{=} \Pi_t^{N,0}, \qquad \overline{\Pi}_\infty^{N,0} \stackrel{\text{(d)}}{=} \frac{1}{N} \sum_{y \in \pi_\infty^0} \delta_{y/\sqrt{N}}, \tag{6.3}$$

where $\Pi_t^{N,0}$ was introduced in Section 5 and π_∞^0 in Section 1. (Since $d \geq 3$, $m_N = N$ here.)

Theorem 3 is equivalent to

$$\lim_{N\to\infty} E\Big[F\Big(\overline{\Pi}_\infty^{N,0}\Big)\Big] = E[F(\mathscr{T}_\infty)] \tag{6.4}$$

for all $F \in C_b(\mathscr{M}(\mathbb{R}^d))$. Since

$$\lim_{t\to\infty} E[F(\mathscr{T}_t)] = E[F(\mathscr{T}_\infty)]$$

is an immediate consequence of Lemma 3 and 3.12, and since by the limit (5.1), $\lim_{N\to\infty} E[F(\overline{\Pi}_t^{N,0})] = E[F(\mathscr{T}_t)]$ holds, it suffices to show that

$$\lim_{t\to\infty} \sup_N \Big| E\Big[F\big(\overline{\Pi}_t^{N,0}\big)\Big] - E\Big[F\big(\overline{\Pi}_\infty^{N,0}\big)\Big]\Big| = 0. \tag{6.5}$$

It is simple to check that the topology of vague convergence on $\mathscr{M}(\mathbb{R}^d)$ is generated by a metric given by a weighted sum of differences as in (4.11), but where the functions f also have compact support. By reasoning analogous to that in the paragraph before the proof of Theorem 4, it suffices to consider, for each compact set $\Gamma \subset \mathbb{R}^d$, those F satisfying

$$|F(\mu) - F(\nu)| \leq \sup_{f \in B_L^\Gamma(\mathbb{R}^d)} |\mu(f) - \nu(f)|,$$

where $B_L^\Gamma(\mathbb{R}^d)$ is the collection of nonnegative, continuous functions f on $\mathbb{R}^d$ which have support in Γ, and are bounded above by 1. For such f,

$$\begin{aligned}\big|\overline{\Pi}_t^{N,0}(f) - \overline{\Pi}_\infty^{N,0}(f)\big| &\leq N^{-1} \sum_{x \in \Gamma \cap \mathbf{S_N}} |1_{\bar{\pi}_t^{N,0}}(x) - 1_{\bar{\pi}_\infty^{N,0}}(x)| \\ &= N^{-1} \sum_{x \in \Gamma \cap \mathbf{S_N}} 1\big\{t < \tau^N(x) < \infty\big\}.\end{aligned}$$

Therefore,

$$\Big| E\Big[F\big(\overline{\Pi}_t^{N,0}\big)\Big] - E\Big[F\big(\overline{\Pi}_\infty^{N,0}\big)\Big]\Big| \leq N^{-1} |\Gamma \cap \mathbf{S_N}| \sup_{x \in \Gamma \cap \mathbf{S_N}} P\big(t < \tau^N(x) < \infty\big). \tag{6.6}$$

To estimate this last probability, we note that $\tau^N(x)$ is the time at which the rate-$2N$ random walk $W_s^{N,x} - W_s^{N,0}$ first hits 0. Therefore, by a standard random walk calculation and the local central limit theorem (see, e.g., the appendix in [7]), for t bounded away from 0,

$$\begin{aligned} P(t < \tau^N(x) < \infty) &\le 2N \int_t^\infty P(W_s^{N,x} = 0)\,ds \le CN \int_t^\infty (sN)^{-d/2}\,ds \\ &= \frac{2C}{d-2}(tN)^{1-d/2} \end{aligned} \tag{6.7}$$

for some finite constant C. Since Γ is compact, $|\Gamma \cap \mathbf{S_N}| \le C' N^{d/2}$ for some C'. On account of this, (6.6) and (6.7), for appropriate C'' and all $N \ge 1$,

$$\left| E\Big[F\big(\overline{\Pi}_t^{N,0}\big)\Big] - E\Big[F\big(\overline{\Pi}_\infty^{N,0}\big)\Big]\right| \le C'' t^{1-d/2}.$$

This proves (6.5). □

7. Weak convergence of random sets. In this section, we demonstrate the convergence of the random sets in Theorems 1′ and 2′. These results are modifications of Theorems 1 and 2, which demonstrate convergence for the corresponding measures. The main step will be given by Lemma 8, which, in essence, states that off a set of small probability, sites in ξ_t^0 will always be near a significant concentration of other sites in ξ_t^0. This prevents the limits in Theorems 1′ and 2′, under the Hausdorff metric, from being larger than the corresponding limits in Theorems 1 and 2. Throughout this section and the next one, condition (1.2) will be assumed.

We consider the family $\{(W_t^{N,x})_{t\ge 0},\ x \in \mathbb{Z}^d\}$ of coalescing random walks used in the previous section, but now with $N = 1$, and denote the family by $\{(W_t^x)_{t\ge 0}, x \in \mathbb{Z}^d\}$. Recall that these are rate-1 random walks with jump kernel $p(x, y)$; the corresponding transition kernels will be denoted by $q_t(x, y)$. For every $t \ge 0$ and $x \in \mathbb{Z}^d$, set

$$\mathscr{V}_t^y = \{x\colon W_t^x = y\}$$

and $\mathscr{V}_t = \mathscr{V}_t^0$. We denote by P_t^* the conditional probability

$$P_t^*(\,\cdot\,) = P\big(\cdot \mid \mathscr{V}_t \ne \varnothing\big).$$

By (2.2), the random sets ξ_t^0 and $\mathscr{V}_t$ have the same distribution. In particular, $p_t = P(\xi_t^0 \ne \varnothing) = P(\mathscr{V}_t \ne \varnothing)$, and so Theorem 1′ is equivalent to the following proposition.

PROPOSITION 1. *The law of* $\frac{1}{\sqrt{t}}\mathscr{V}_t$ *under* P_t^* *converges weakly to the law of* $\operatorname{supp}\mu$ *under* $\widehat{R}_1(0, d\mu)$.

The following lemma will be employed in Lemma 8, which will then be used to demonstrate Proposition 1.

LEMMA 7. *There exist positive constants C and C' such that, for every $t > 1$ and $A > 0$,*

$$P_t^*\left(\sup_{x\in\mathscr{V}_t}|x| > A\sqrt{t}\right) \le C\exp(-C'A). \tag{7.1}$$

PROOF. For $A > 0$ and $n \ge 1$, set $A_n = \frac{1}{12}A\sum_{k=1}^{n} 2^{-k/4}$ and set $A_0 = 0$. Also, for $t > 1$, denote by $N = N(t)$ the first integer such that $2^{-N}t < 1$. It is easy to check that the event on the left side of (7.1) can only occur if, for some $x \in \mathscr{V}_t$ with $|x| > A\sqrt{t}$, one of the following three events occurs:

(a) $|W^x_{t/2^N}| \le A_N\sqrt{t}$.

(b) $|W^x_{t/2^{n+1}}| > A_{n+1}\sqrt{t}$ and $|W^x_{t/2^n}| \le A_n\sqrt{t}$ for some $n = 1, \dots, N-1$.

(c) $|W^x_{t/2}| > \frac{2^{-1/4}}{12}A\sqrt{t}$.

We will obtain upper bounds on the probabilities of each of these three possibilities. In each case we will use

$$P\big(|W_t| > A\sqrt{t}\big) \le c_1\exp(-c_2A), \tag{7.2}$$

where $c_1 > 0$ and $c_2 > 0$ do not depend on $t > 1/2$ and $A > 0$; this inequality is a straightforward consequence of the assumption (1.2).

We first consider (c). Set $A' = \frac{2^{-1/4}}{12}A$. For every $t > 1$,

$$P\big(\exists\ x \in \mathscr{V}_t\colon |W^x_{t/2}| > A'\sqrt{t}\big) \le E\Bigg[\sum_{|y|>A'\sqrt{t}} 1_{\{\mathscr{V}^y_{t/2}\neq\varnothing,\,\mathscr{V}^y_{t/2}\subset\mathscr{V}_t\}}\Bigg].$$

[When interpreted in terms of the voter model over $[0, t]$, the event in the indicator function on the right side above is the event that the opinion at $(t/2, y)$ is "descended" from that at $(0, 0)$, and itself has "descendants" at time t.] By using the Markov property at time $t/2$, this expectation equals $\sum_{|y|>A'\sqrt{t}} p_{t/2}\, q_{t/2}(y, 0)$. It follows, using (7.2), that

$$P\big(\exists\ x \in \mathscr{V}_t\colon |W^x_{t/2}| > A'\sqrt{t}\big) \le c_1\, p_{t/2}\, \exp(-c_2A'). \tag{7.3}$$

We next consider (b). For every $n = 1, \dots, N-1$,

$$\begin{aligned}
&P\Big(\exists\ x \in \mathscr{V}_t\colon\ |W^x_{t/2^{n+1}}| > A_{n+1}\sqrt{t} \text{ and } |W^x_{t/2^n}| \le A_n\sqrt{t}\Big)\\
&\quad\le \sum_{|y|>A_{n+1}\sqrt{t}}\ \sum_{|z|\le A_n\sqrt{t}} P\big(\exists\ x\colon W^x_{t/2^{n+1}} = y,\ W^x_{t/2^n} = z,\ W^x_t = 0\big)\\
&\quad= \sum_{|y|>A_{n+1}\sqrt{t}}\ \sum_{|z|\le A_n\sqrt{t}} p_{t/2^{n+1}}\, q_{t/2^{n+1}}(y, z)\, q_{t-(t/2^n)}(z, 0) \qquad (7.4)\\
&\quad\le p_{t/2^{n+1}}\, P\Big(|W_{t/2^{n+1}}| \ge (A_{n+1} - A_n)\sqrt{t}\Big)\\
&\quad\le c_1\, p_{t/2^{n+1}}\, \exp\left(-\frac{2^{(n+1)/4}}{12}\, c_2A\right).
\end{aligned}$$

The reasoning for (a) is similar. One has

$$
\begin{aligned}
P\Big(\exists\, x \in \mathscr{V}_t\colon |W^x_{t/2^N}| \le A_N\sqrt{t} \text{ and } |x| > A\sqrt{t}\Big)
&\le \sum_{|x|>A\sqrt{t}} \sum_{|y|\le A_N\sqrt{t}} P(W^x_{t/2^N} = y,\ W^x_t = 0)\\
&= \sum_{|x|>A\sqrt{t}} \sum_{|y|\le A_N\sqrt{t}} q_{t/2^N}(x, y)\, q_{t-(t/2^N)}(y, 0) \qquad (7.5)\\
&\le c_1 \exp\Big(-2^{N/2} c_2 A/2\Big)\\
&\le c_1 \exp\Big(-c_2\sqrt{t}A/2\Big),
\end{aligned}
$$

since $A_N \le A/2$.

Putting together (7.3), (7.4) and (7.5), we arrive at

$$
\begin{aligned}
P\Big(\sup_{x\in\mathscr{V}_t} |x| > A\sqrt{t}\Big) \le{}& c_1\, p_{t/2} \exp\Big(-\frac{2^{-1/4}}{12} c_2 A\Big)\\
&+ c_1 \sum_{n=1}^{N-1} p_{t/2^{n+1}} \exp\Big(-\frac{2^{(n+1)/4}}{12} c_2 A\Big)\\
&+ c_1 \exp\big(-c_2 A\sqrt{t}/2\big).
\end{aligned}
$$

The inequality (7.1), for $A \ge 1$, follows from this bound and (1.5). Increasing C by the factor $e^{C'}$ implies (7.1) all $A > 0$. □

For $a \in \mathbb{R}^d$ and $r > 0$, we denote by $B(a, r)$ the open ball of radius r centered at a. Lemma 8 shows that, with high probability, there are many other points of $\mathscr{V}_t$ near *every* point of $\mathscr{V}_t$. This result provides the main step in the proofs of Propositions 1 and 2 and of Theorem 5 at the end of the section.

LEMMA 8. *Let $\rho > 0$ and $\eta > 0$. For small enough $\delta > 0$ and large enough t,*

$$
P^*_t\Big(\exists\, x \in \mathscr{V}_t\colon |\mathscr{V}_t \cap B(x, \eta\sqrt{t})| < \delta m_t\Big) < \rho. \qquad (7.6)
$$

PROOF. Inequality (7.6) can be motivated in terms of the voter model over $[0, t]$. We will argue that, except on a set of small probability, (a) all "ancestors" at time $(1-\varepsilon)t$, where $\varepsilon > 0$ is fixed, are "close" to their "descendants" at time t, and (b) all such ancestors have at least of order of magnitude m_t descendants. Part (a) will follow from Lemma 7 and is given in (7.7); part (b) is given in (7.8).

We first consider (a). For every $\varepsilon \in (0, 1]$, let

$$
\mathscr{W}_{\varepsilon, t} = \big\{y \in \mathbb{Z}^d\colon \mathscr{V}^y_{\varepsilon t} \ne \varnothing \text{ and } \mathscr{V}^y_{\varepsilon t} \subset \mathscr{V}_t\big\}.
$$

[For the voter model, this is the set of all descendants at time $(1-\varepsilon)t$, of the opinion at the origin at time 0, that themselves have descendants at time t. Recall that time for the voter model runs backwards relative to the random walks W^x_t.] By applying the Markov property at time εt, and then Lemma 7, we get, for every $\gamma > 0$ and $\varepsilon \in (0, 1/2)$,

$$\begin{aligned} P\Big(\exists\ y \in \mathscr{W}_{\varepsilon,t}\colon \mathscr{V}^y_{\varepsilon t} \not\subset B\big(y, \gamma\sqrt{t}\big)\Big) &\le \textstyle\sum_{y\in\mathbb{Z}^d} P\Big(\mathscr{V}^y_{\varepsilon t} \not\subset B\big(y, \gamma\sqrt{t}\big)\Big)\, q_{(1-\varepsilon)t}(y, 0) \\ &\le C\, p_{\varepsilon t} \exp\Big(-C' \frac{\gamma}{\sqrt{\varepsilon}}\Big), \end{aligned} \tag{7.7}$$

provided that t is sufficiently large. The constants C and C', from Lemma 7, do not depend on ε.

Recall from (1.6) that the law of $p_t|\mathscr{V}_t|$, under P^*_t, converges, as $t \to \infty$, to an exponential distribution with parameter 1, that is, for any $\alpha > 0$,

$$\lim_{t\to\infty} P^*_t\big(p_t|\mathscr{V}_t| \le \alpha\big) = 1 - e^{-\alpha} < \alpha.$$

Using the same decomposition as in (7.7), we have, for given $\varepsilon \in (0, 1/2)$ and t sufficiently large,

$$\begin{aligned} P\big(\exists\ y \in \mathscr{W}_{\varepsilon,t}\colon |\mathscr{V}^y_{\varepsilon t}| \le \alpha\, p^{-1}_{\varepsilon t}\big) &\le \sum_{y\in\mathbb{Z}^d} P\big(0 < |\mathscr{V}^y_{\varepsilon t}| \le \alpha\, p^{-1}_{\varepsilon t}\big)\, q_{(1-\varepsilon)t}(y, 0) \\ &= p_{\varepsilon t}\, P^*_{\varepsilon t}\big(|\mathscr{V}_{\varepsilon t}| \le \alpha\, p^{-1}_{\varepsilon t}\big) \\ &\le p_{\varepsilon t}\alpha. \end{aligned} \tag{7.8}$$

By combining (7.7) and (7.8), we see that, for any fixed $\gamma > 0$, $\alpha > 0$ and $\varepsilon \in (0, 1/2)$, and large t,

$$\begin{aligned} &P^*_t\big(\exists\ y \in \mathscr{W}_{\varepsilon,t}\colon \mathscr{V}^y_{\varepsilon t} \not\subset B(y, \gamma\sqrt{t}) \text{ or } |\mathscr{V}^y_{\varepsilon t}| \le \alpha\, p^{-1}_{\varepsilon t}\big) \\ &\quad\le \frac{p_{\varepsilon t}}{p_t}\Big(C \exp\Big(-C' \frac{\gamma}{\sqrt{\varepsilon}}\Big) + \alpha\Big). \end{aligned} \tag{7.9}$$

Last, we consider the behavior of $\mathscr{V}_t$ on the complement of the event in (7.9) and set

$$H = \big\{\forall y \in \mathscr{W}_{\varepsilon,t}, \mathscr{V}^y_{\varepsilon t} \subset B\big(y, \gamma\sqrt{t}\big) \text{ and } |\mathscr{V}^y_{\varepsilon t}| > \alpha\, p^{-1}_{\varepsilon t}\big\}.$$

For any given $x \in \mathscr{V}_t$, set $y = W^x_{\varepsilon t} \in \mathscr{W}_{\varepsilon,t}$. Then, on H,

$$\begin{aligned} |\mathscr{V}_t \cap B(x, 2\gamma\sqrt{t})| &\ge |\mathscr{V}_t \cap B(y, \gamma\sqrt{t})| \\ &\ge |\mathscr{V}^y_{\varepsilon t}| \ge \alpha\, p^{-1}_{\varepsilon t} \ge \varepsilon\alpha\beta_d m_t/2 \end{aligned} \tag{7.10}$$

for each $x \in \mathscr{V}_t$, where the first bound follows from $|y - x| \le \gamma\sqrt{t}$, and the last bound holds for t large enough because of (1.5).

If one sets $\eta = 2\gamma$ and $\delta = \varepsilon\alpha\beta_d/2$, the inner inequality in (7.6) does not hold on H, and so the left side of (7.6) is bounded above by $P(H^c)$. Moreover, if one chooses $\varepsilon > 0$ and $\alpha > 0$ small enough so that

$$\frac{2}{\varepsilon}\Big(C \exp\Big(-C' \frac{\gamma}{\sqrt{\varepsilon}}\Big) + \alpha\Big) < \rho,$$

then $P(H^c) < \rho$ for large t, because of (7.9) and (1.5). This implies (7.6). $\square$

PROOF OF PROPOSITION 1. It is enough to show convergence along each sequence $t_n \uparrow \infty$. For every $t > 0$, let Z_t be the random measure defined by

$$Z_t = \frac{1}{m_t} \sum_{y \in \mathscr{Y}_t} \delta_{y/\sqrt{t}}.$$

By Theorem 1, the law of Z_t under P_t^* converges weakly to $\widehat{R}_1(0, \cdot)$. So, by the Skorokhod representation theorem, there exist random measures $\widetilde{Z}_{t_n}$, defined on the same probability space, such that for every n, $\widetilde{Z}_{t_n}$ has the law of Z_{t_n} under $P_{t_n}^*$, and

$$\widetilde{Z}_{t_n} \longrightarrow \widetilde{Z}_\infty \qquad \text{a.s.}, \tag{7.11}$$

where $\widetilde{Z}_\infty$ has distribution $\widehat{R}_1(0, \cdot)$.

Recall that the Hausdorff metric on nonempty compact subsets of $\mathbb{R}^d$ is defined by $d_0(K, K') = d_1(K, K') + d_1(K', K)$, where $d_1(K, K') = \inf\{\varepsilon > 0\colon K \subset K'_\varepsilon\}$ and K'_ε denotes the closed ε-enlargement of K'. To show Proposition 1, it is enough to verify that

$$d_0\big(\operatorname{supp} \widetilde{Z}_{t_n}, \operatorname{supp} \widetilde{Z}_\infty\big) \longrightarrow 0$$

in probability as $n \to \infty$. It is well known, and easy to prove, that (7.11) implies

$$d_1(\operatorname{supp} \widetilde{Z}_\infty, \operatorname{supp} \widetilde{Z}_{t_n}) \longrightarrow 0 \quad \text{a.s.}$$

(In order for $\widetilde{Z}_{t_n}$, as $n \to \infty$, to contribute mass arbitrarily close to some point z, $\widetilde{Z}_{t_n}$ must also contain sites which are close.) Thus, the problem is to prove that

$$d_1\big(\operatorname{supp} \widetilde{Z}_{t_n}, \operatorname{supp} \widetilde{Z}_\infty\big) \longrightarrow 0 \tag{7.12}$$

in probability.

Fix $\alpha > 0$ and $\gamma > 0$. From Lemma 8 and the definition of Z_t, we can choose $\delta > 0$ small enough so that for every t large enough,

$$P_t^*\Big(\exists\, z \in \operatorname{supp} Z_t\colon Z_t\Big(B\Big(z, \frac{\alpha}{2}\Big)\Big) < \delta\Big) < \frac{\gamma}{2}. \tag{7.13}$$

From the definition of d_1,

$$\begin{aligned} &P\Big(d_1\big(\operatorname{supp} \widetilde{Z}_{t_n}, \operatorname{supp} \widetilde{Z}_\infty\big) > \alpha\Big) \\ &\qquad = P\Big(\exists\, z \in \operatorname{supp} \widetilde{Z}_{t_n}\colon \operatorname{dist}\big(z, \operatorname{supp} \widetilde{Z}_\infty\big) > \alpha\Big). \end{aligned}$$

Using (7.13) and the fact that $\widetilde{Z}_{t_n}$ has the law of Z_{t_n} under $P_{t_n}^*$, we see that, for n large enough, the previous quantity is bounded above by

$$\frac{\gamma}{2} + P\Big(\exists\, z \in \mathbb{R}^d\colon \widetilde{Z}_{t_n}\Big(B\Big(z, \frac{\alpha}{2}\Big)\Big) \ge \delta \text{ and } \widetilde{Z}_\infty\big(B(z, \alpha)\big) = 0\Big). \tag{7.14}$$

Recall the definition (4.11) of the metric d inducing the weak topology on $\mathscr{M}_f(\mathbb{R}^d)$, and note that for the function $f(y) = (\alpha - |z-y|)^+$, $|\widetilde{Z}_{t_n}(f) - \widetilde{Z}_\infty(f)| \geq \alpha\delta/2$ on the event in (7.14). It therefore follows from (7.14) that, for large n,

$$P\Big(d_1\big(\operatorname{supp}\widetilde{Z}_{t_n}, \operatorname{supp}\widetilde{Z}_\infty\big) > \alpha\Big) \leq \frac{\gamma}{2} + P\Big(d\big(\widetilde{Z}_{t_n}, \widetilde{Z}_\infty\big) \geq \alpha\delta/2\Big).$$

By (7.11), this is bounded above by γ for n large enough. Since γ can be chosen arbitrarily close to 0, this completes the proof. $\square$

We now demonstrate Theorem 2′. The set $\bar{\pi}_t^0 = \bar{\pi}_t^{1,0}$, defined in Section 6, has the same distribution as π_t^0. It therefore suffices to prove the following.

PROPOSITION 2. *The random sets* $\frac{1}{\sqrt{t}}\bar{\pi}_t^0$ *converge in distribution to* $\operatorname{supp}\mathscr{T}_1$.

PROOF. We wish to show that the following analog of Lemma 8 holds: for every $\rho > 0$ and $\eta > 0$, if $\delta > 0$ is chosen small enough and t large enough,

$$P\Big(\exists\, x \in \bar{\pi}_t^0\colon |\bar{\pi}_t^0 \cap B(x, \eta\sqrt{t})| < \delta m_t\Big) < \rho. \tag{7.15}$$

Once one has shown (7.15), the argument is the same as that given in the proof of Proposition 1, which we therefore omit.

In order to show (7.15), first recall from (1.3)–(1.5), that $p_t|\bar{\pi}_t^0|$ converges in distribution as $t \to \infty$. We can therefore choose $M > 0$ large enough so that for every $t > 0$,

$$P\big(p_t\,|\bar{\pi}_t^0| > M\big) < \frac{\rho}{2}. \tag{7.16}$$

Let $\mathscr{A}$ denote the collection of finite subsets of $\mathbb{Z}^d$. For any $z \in \mathbb{Z}^d$ and $A \in \mathscr{A}$ with $0 \in A$, $\{\bar{\pi}_t^0 = A, W_t^0 = z\} = \{\mathscr{V}_t^z = A\}$. Also, let $h(A) = 1$ for those sets A with $|A| \leq M p_t^{-1}$ and such that $|A \cap B(x, \eta\sqrt{t})| < \delta m_t$ for some $x \in A$, and set $h(A) = 0$ otherwise. After a simple decomposition, this implies

$$\begin{aligned} &P\Big(|\bar{\pi}_t^0| \leq M p_t^{-1} \text{ and } \exists\, x \in \bar{\pi}_t^0\colon |\bar{\pi}_t^0 \cap B\big(x, \eta\sqrt{t}\big)| < \delta m_t\Big) \\ &\quad= \sum_{z\in\mathbb{Z}^d} \sum_{A\in\mathscr{A}\colon\, 0\in A} P\big(\bar{\pi}_t^0 = A, W_t^0 = z\big) h(A) \\ &\quad= \sum_{z\in\mathbb{Z}^d} \sum_{A\in\mathscr{A}\colon\, 0\in A} P(\mathscr{V}_t^z = A) h(A). \end{aligned}$$

Since $P(\mathscr{V}_t^z = A) = P(\mathscr{V}_t = A - z)$, and $h(A) = h(A+z)$, by changing variables and interchanging the order of summation, we have

$$\begin{aligned} \sum_{z\in\mathbb{Z}^d} \sum_{A\in\mathscr{A}\colon\, 0\in A} P(\mathscr{V}_t^z = A) h(A) &= \sum_{A\in\mathscr{A}} \sum_{z\in A} P(\mathscr{V}_t = A) h(A) \\ &= \sum_{A\in\mathscr{A}} P(\mathscr{V}_t = A)|A| h(A). \end{aligned}$$

This is at most $MP_t^*(\exists\, x \in \mathscr{V}_t\colon |\mathscr{V}_t \cap B(x, \eta\sqrt{t})| < \delta m_t)$, which, by Lemma 8, is at most $\rho/2$ for large t. Putting things together, it follows that

$$P\Big(|\bar{\pi}_t^0| \le M p_t^{-1} \text{ and } \exists\, x \in \bar{\pi}_t^0\colon |\bar{\pi}_t^0 \cap B(x, \eta\sqrt{t})| < \delta m_t\Big) \le \rho/2 \tag{7.17}$$

for large t. Combining (7.16) and (7.17), we obtain (7.15). □

Let X_t^N be defined as above (1.9), and assume that $X_0^N \to X_0 \in \mathscr{M}_f(\mathbb{R}^d)$ as $N \to \infty$. In (1.9), the result $(X_t^N)_{t\ge 0} \Rightarrow (X_t)_{t\ge 0}$, where X_t is super-Brownian motion with branching rate $2\beta_d$ and diffusion coefficient σ^2, was quoted from [7]. The ideas from Section 7 can also be used to give a "set version" of this result.

THEOREM 5. *The set-valued process $(\xi_t^N)_{t>0}$ converges in distribution to $(\operatorname{supp} X_t)_{t>0}$, in the sense of weak convergence of finite-dimensional marginals.*

We exclude $t = 0$ in Theorem 5, since our assumptions do not imply the convergence of the sets ξ_0^N, and furthermore, $\operatorname{supp} X_0$ need not be compact. [For $t > 0$, $\operatorname{supp} X_t$ is a.s. compact (see Section 9.3 in [9]).]

PROOF. As in Proposition 2, it suffices to demonstrate the analog of Lemma 8 for the random sets ξ_t^N, for each fixed $t > 0$. Namely, we wish to verify, for each choice of $\rho > 0$ and $\eta > 0$, that for $\delta > 0$ sufficiently small and N sufficiently large,

$$P\Big(\exists\, x \in \xi_t^N\colon |\xi_t^N \cap B(x, \eta)| < \delta m_N\Big) < \rho. \tag{7.18}$$

The remainder of the argument is then the same as in the proof of Proposition 1.

The left side of (7.18) is bounded above by

$$\begin{aligned}
&\sum_{y\in\xi_0^N} P\Big(\exists\, x \in \xi_t^{N,y}\colon |\xi_t^{N,y} \cap B(x, \eta)| < \delta m_N\Big)\\
&\qquad = p_{Nt}\,|\xi_0^N|\, P\Big(\exists\, x \in \xi_t^{N,0}\colon |\xi_t^{N,0} \cap B(x, \eta)| < \delta m_N \mid \xi_t^{N,0} \ne \varnothing\Big).
\end{aligned}$$

The assumption $X_0^N \to X_0$ implies that $p_{Nt}\, |\xi_0^N|$ remains bounded, in probability, as $N \to \infty$. Since $\xi_t^{N,0}$ and $\frac{1}{\sqrt{N}}\mathscr{V}_{Nt}$ have the same distribution, (7.18) follows from Lemma 8. □

8. A related diffusion equation. Proposition 1 can be used to answer questions of the following type. Let A be an open subset in $\mathbb{R}^d$. What is the limiting behavior of the probability that the voter model, starting from a single 1 at the site 0, intersects $\sqrt{t}A$ at time t? One can also phrase the problem in terms of a system of coalescing random walks starting at every point of $\sqrt{t}A \cap \mathbb{Z}^d$: What is the limiting behavior of the probability that one of these walks is at the origin at time t?

Let A be an open subset of $\mathbb{R}^d$. We say that A satisfies the interior cone condition if, for every point $z \in \partial A$, there is an open cone with vertex z which is contained in A in the neighborhood of z.

THEOREM 6. *Suppose that A satisfies the interior cone condition. Then,*

$$\begin{aligned}\lim_{t\to\infty} p_t^{-1} P(\xi_t^0 \cap \sqrt{t}A \neq \varnothing) &= \lim_{t\to\infty} p_t^{-1} P(\mathscr{Y}_t \cap \sqrt{t}A \neq \varnothing)\\ &= \int_{\{\operatorname{supp}\mu\cap A\neq\varnothing\}} \widehat{R}_1(0, d\mu).\end{aligned} \tag{8.1}$$

This limit equals $u_1(0)$, where the function $(u_t(x),\ t > 0,\ x \in \mathbb{R}^d)$ is the unique nonnegative solution of the problem

$$\begin{aligned}&\frac{\partial u}{\partial t} = \frac{\sigma^2}{2}\Delta u - u^2, \qquad (t, x) \in (0, \infty) \times \mathbb{R}^d,\\ &u_0(x) = +\infty, \qquad x \in A,\\ &u_0(x) = 0, \qquad x \in \mathbb{R}^d \setminus \overline{A},\end{aligned} \tag{8.2}$$

where $\overline{A}$ denotes the closure of A.

PROOF. For every $t > 0$ and $x \in \mathbb{R}^d$, set

$$\begin{aligned}v_t(x) &= \int_{\{\operatorname{supp}\mu\cap A\neq\varnothing\}} R_t(x, d\mu),\\ \bar{v}_t(x) &= \int_{\{\operatorname{supp}\mu\cap \overline{A}\neq\varnothing\}} R_t(x, d\mu).\end{aligned}$$

By known connections between superprocesses and partial differential equations (see [12]), the function $v_t(x)$ is the minimal nonnegative solution of the problem

$$\begin{aligned}&\frac{\partial v}{\partial t} = \frac{\sigma^2}{2}\Delta v - \beta_d\, v^2, \qquad (t, x) \in (0, \infty) \times \mathbb{R}^d,\\ &v_0(x) = +\infty, \qquad x \in A.\end{aligned}$$

Similarly, $\bar{v}_t(x)$ is the maximal nonnegative solution of the problem

$$\begin{aligned}&\frac{\partial \bar{v}}{\partial t} = \frac{\sigma^2}{2}\Delta \bar{v} - \beta_d\, \bar{v}^2, \qquad (t, x) \in (0, \infty) \times \mathbb{R}^d,\\ &\bar{v}_0(x) = 0, \qquad x \in \mathbb{R}^d \setminus \overline{A}.\end{aligned}$$

From arguments similar to the proof of Theorem 7.1 in [1], one easily sees that the interior cone condition implies $v_t(x) = \bar{v}_t(x)$ for every t and x. It follows that the function $v_t(x)$ is the unique nonnegative solution of (8.2), with u^2 replaced by $\beta_d u^2$. Obviously, $u_t(x) = \beta_d v_t(x)$ is then the unique nonnegative solution of (8.2).

We now show (8.1). Observe that the set of all compact subsets K of $\mathbb{R}^d$, with $K \cap A \neq \varnothing$, is open with respect to the Hausdorff metric. It follows from Proposition 1 that

$$\liminf_{t\to\infty} P_t^*(\mathscr{V}_t \cap \sqrt{t}A \neq \varnothing) \geq \int_{\{\operatorname{supp}\mu\cap A\neq\varnothing\}} \widehat{R}_1(0, d\mu) = \beta_d\, v_1(0).$$

Similarly, since the set of all compact sets K such that $K \cap \overline{A} \neq \varnothing$ is closed,

$$\limsup_{t\to\infty} P_t^*(\mathscr{V}_t \cap \sqrt{t}\overline{A} \neq \varnothing) \leq \int_{\{\operatorname{supp}\mu\cap \overline{A}\neq\varnothing\}} \widehat{R}_1(0, d\mu) = \beta_d \bar{v}_1(0).$$

The equality $v_1(0) = \bar{v}_1(0)$ then gives (8.1). $\square$

It is interesting to compare Theorem 6 with Sznitman's results [24] about systems of annihilating Brownian spheres in $\mathbb{R}^d$. Sznitman studies the limiting behavior of such a system when the radius of the spheres tends to 0 and the initial number of particles goes to ∞. The limiting density of particles is then given as a solution of the same equation as in Theorem 6, but with a different constant in the forcing term; the initial value also differs because Sznitman starts with a given initial density of particles. Such a connection is not too surprising on account of a result in [3], where it is shown that the limiting density of particles, except for a constant factor 2, is the same for systems of coalescing and annihilating random walks.

REFERENCES

[1] ABRAHAM, R. and LE GALL, J. F. (1994). Sur la mesure de sortie du super-mouvement brownien. *Probab. Theory Related Fields* **99** 251–275.

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

[3] ARRATIA, R. (1981). Limiting point processes for rescalings of coalescing and annihilating random walks on $\mathbb{Z}^d$. *Ann. Probab.* **9** 909–936.

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

[5] CLIFFORD, P. and SUDBURRY, A. (1973). A model for spatial conflict. *Biometrika* **60** 581–588.

[6] COX, J. T. and DURRETT, R. (1995). Hybrid zones and voter model interfaces. *Bernoulli* **1** 343–370.

[7] COX, J. T., DURRETT, R. and PERKINS, E. (2000). Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* **28** 185–234.

[8] DAWSON, D. A. (1975). Stochastic evolution equations and related measures processes. *J. Multivariate Anal.* **3** 1–52.

[9] DAWSON, D. A. (1993). Measure-valued Markov processes. *École d'Été de Probabilités de Saint Flour XXI*. *Lecture Notes in Math.* **1541** 1–260. Springer, Berlin.

[10] DAWSON, D. A. and PERKINS, E. (1991). Historical processes. *Mem. Amer. Math. Soc.* **454** 1–179.

[11] DURRETT, R. (1996). *Stochastic Spatial Models.* PCMI Lecture Notes, IAS, Princeton.

[12] DYNKIN, E. B. (1993). Superprocesses and partial differential equations. *Ann. Probab.* **21** 1185–1262.

[13] EL KAROUI, N. and ROELLY, S. (1991). Propriétés de martingales, explosion et représentation de Lévy-Khintchine d'une classe de processus de branchement à valeurs mesures. *Stochastic Process. Appl.* **38** 239–266.

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

[15] GRIFFEATH, D. (1979). *Additive and Cancellative Interacting Particle Systems.* Springer, New York.

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

[17] LE GALL, J.-F. (1991). Brownian excursions, trees and measure-valued branching processes. *Ann. Probab.* **19** 1399–1439.

[18] LE GALL, J.-F. (1999). *Spatial Branching Processes, Random Snakes, And Partial Differential Equations.* Birkhäuser, Boston.

[19] LE GALL, J.-F. and PERKINS, E. (1995). The Hausdorff measure of the support of two-dimensional super-Brownian motion. *Ann. Probab.* **23** 1719–1747.

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

[21] PERKINS, E. (1999). Measure-valued processes and interactions. *École d'Été de Probabilités de Saint Flour*. *Lecture Notes in Math.* Springer, Berlin. To appear.

[22] SAWYER, S. (1979). A limit theorem for patch sizes in a selectively-neutral migration model. *J. Appl. Probab.* **16** 482–495.

[23] SPITZER, F. L. (1976). *Principles of Random Walk.* Springer, New York.

[24] SZNITMAN, A. S. (1988). Propagation of chaos for a system of annihilating Brownian spheres. *Comm. Pure Appl. Math.* **60** 663–690.

[25] WATANABE, S. (1968). A limit theorem of branching processes and continuous state branching. *J. Math. Kyoto U.* **8** 141–167.

[26] ZÄHLE, I. (1999). Renormalization of the voter model in equilibrium. Preprint.

SCHOOL OF MATHEMATICS
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA 55455
E-MAIL: bramson@math.umn.edu

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

LABORATOIRE DE MATHEMATIQUES
DE L'ECOLE NORMALE SUPERIÉURE
45 RUE D'ULM 75320 PARIS CEDEX 05
FRANCE
E-MAIL: legall@dmi.ens.fr