*The Annals of Probability*
2000, Vol. 28, No. 1, 185–234

# RESCALED VOTER MODELS CONVERGE TO SUPER-BROWNIAN MOTION

BY J. THEODORE COX,[1] RICHARD DURRETT[2] AND EDWIN A. PERKINS[3]

*Syracuse University, Cornell University and University of British Columbia*

We show that a sequence of voter models, suitably rescaled in space and time, converges weakly to super-Brownian motion. The result includes both nearest neighbor and longer range voter models and complements a limit theorem of Mueller and Tribe in one dimension.

**1. Introduction.** Super-Brownian motion and its close relatives have been studied by many authors [see Dawson (1993), Dynkin (1994), Dawson and Perkins (1991, 1998), Le Gall (1994), and the references therein]. These processes originally arose as weak limits of rescaled branching random walks [see Watanabe (1968) and Theorem 1.0 below]. It has recently been shown that a broad range of more complex interacting spatial models, when suitably rescaled, converge to super-Brownian motion (or a process closely related to it). Examples include rescaled limits of random trees [Derbez and Slade (1997)] and limits of long-range contact processes in dimensions $d \geq 2$ [Durrett and Perkins (1999)]. Ongoing work of Derbez, van der Hofstad and Slade (1998) will almost certainly add oriented percolation at the critical probability in high dimensions to this list.

Our goal in this work is to show that rescaled voter models (to be defined in a moment) in dimensions two or more converge to super-Brownian motion. This convergence will be established in a variety of scenarios. First in the case of long-range interactions and then, more surprisingly, for nearest neighbor interactions. To explain this in more detail we have to give the definition of the two main processes we will consider.

1. The *rate one voter model on* $\mathbf{Z}^d$ *with symmetric voting kernel* $p(x, y) = p(y - x) = p(x - y)$ is a Markov process $\xi_t: \mathbf{Z}^d \to \{0, 1\}$, where $\xi_t(x)$ gives the "opinion" (either 0 or 1) of the voter at site $x \in \mathbf{Z}^d$ at time $t$. At exponential times with rate 1, a given individual selects a site at random according to the kernel $p(\cdot)$, and then adopts the opinion of the selected site. Here, the successive waiting times of all the sites, and all the choices according to $p(\cdot)$, are taken to be independent. Clearly, if a site $x$ selects a site $y$ with the same opinion, no change occurs.

Received January 1999.

[1]Supported in part by NSF Grant DMS-96-26675 and by a NSERC Collaborative Grant. Part of the research was done while the author was on sabbatical from Syracuse University and in residence at University of British Columbia.
[2]Supported in part by NSF Grant DMS-96-26201.
[3]Supported in part by a NSERC Research grant and a NSERC Collaborative Grant.
*AMS* 1991 *subject classifications*. Primary 60K35, 60G57; secondary 60F05, 60J80.
*Key words and phrases*. Voter model, super-Brownian motion.

2. The *rate $\gamma$ critical branching random walk on $\mathbf{Z}^d$ with kernel $p(\cdot)$*, is a Markov process $\zeta_t:\mathbf{Z}^d \to \{0, 1, 2, \dots\}$, where $\zeta_t(x)$ gives the number of particles at site $x \in \mathbf{Z}^d$ at time $t$. A particle at $x$ dies with rate $\gamma$, that is, it lives for an exponentially distributed time with mean $1/\gamma$, and with rate $\gamma p(y-x)$ produces a new particle at $y$ for each $y \in \mathbf{Z}^d$. Again, the successive waiting times for each site and each possible pair of sites are independent.

To compare these two processes it is useful to think of the voter model as a branching random walk in which the birth rate from $x$ to $y$ and the death rate at $x$ depend on the number of vacant neighbors. We now reformulate our description of the voter model $\xi_t$ in terms of such a state dependent branching random walk. To do this, we first note that the dynamics of the voter model are clearly equivalent to the following: for each $(x, y) \in \mathbf{Z}^d \times \mathbf{Z}^d$ satisfying $x \neq y$, the individual at $x$ imposes its opinion on the individual at $y$ with rate $p(y-x)$. Now consider sites with opinion 1 as occupied and sites with opinion 0 as unoccupied. With this interpretation the process behaves as follows: at time $t$ each occupied site $x$ produces an offspring at $y$ with rate $p(y-x)\mathbf{1}(\xi_t(y) = 0)$, and dies with rate

$$V_t(x) \equiv \sum_y p(y-x)\mathbf{1}(\xi_t(y) = 0). \tag{1.1}$$

This corresponds to a state dependent branching rate from $x$ to $y$ of $p(y-x)\mathbf{1}(\xi_t(y) = 0)$, which implies a total branching rate from $x$ also equal to $V_t(x)$, the "local density" of 0's near $x$.

To get a measure-valued limit process, we will speed up time by $N$ and scale space down by $\sqrt{N}$. If, after this rescaling, the local densities of 0's at distinct sites separated by a positive distance on the resulting "macroscopic" scale are approximately independent, then we can expect a *mean-field simplification*, and the rescaled voter models should behave like the rescaled branching random walks, but now with $\gamma$ equal to the mean local density of 0's near a typical 1.

Our task in each of the examples we consider will be to justify this mean-field simplification and calculate the effective branching rate $\gamma$. Intuitively, this should be very easy if the kernels $p_N(\cdot)$ become sufficiently spread out as $N \to \infty$. For, in this case, the effective local density of 0's should approach 1, and thus the rescaled voter models should behave like the rescaled branching random walks with $\gamma = 1$. The asymptotic behavior of the latter is well known: it approaches super-Brownian motion (see Theorem 1.0 below). Inspired by our heuristic, we will begin with this case, denoted (1) below.

To prove Theorem 1.1, we need a characterization of the limiting super-Brownian motion. Let $\mathscr{M}_F(\mathbf{R}^d)$ denote the space of finite measures on $\mathbf{R}^d$, endowed with the topology of weak convergence of measures. Let $\Omega_{X,D} = D([0,\infty), \mathscr{M}_F(\mathbf{R}^d))$ be the Skorohod space of cadlag $\mathscr{M}_F(\mathbf{R}^d)$-valued paths, and let $\Omega_{X,C}$ be the space of continuous $\mathscr{M}_F(\mathbf{R}^d)$-valued paths with the topology of uniform convergence on compacts. In either case, $X_t$ will denote the coordinate function, $X_t(\omega) = \omega(t)$. Integration of a function $\phi$ with respect to a measure $\mu$ will be denoted by $\mu(\phi)$. For $1 \le n \le \infty$ let $C_b^n(\mathbf{R}^d)$ be the space of bounded continuous functions whose partial derivatives of order less than $n+1$ are also bounded and continuous, and let $C_0^n(\mathbf{R}^d)$ be the space of those functions in $C_b^n(\mathbf{R}^d)$ with compact support.

An adapted a.s.-continuous $\mathscr{M}_F(\mathbf{R}^d)$-valued process $X_t$, $t \ge 0$ on a complete filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ is said to be a *super-Brownian motion with branching rate b and diffusion coefficient* $\sigma^2 > 0$ *starting at* $X_0 \in \mathscr{M}_F(\mathbf{R}^d)$ if it solves the following martingale problem:

$(\mathrm{MP})_{X_0}^{b,\sigma^2}$ For all $\phi \in C_0^\infty(\mathbf{R}^d)$, $M_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s(\sigma^2 \Delta\phi/2)\, ds$ is a continuous $(\mathscr{F}_t)$-martingale, with $M_0(\phi) = 0$ and square function $\langle M(\phi)\rangle_t = \int_0^t X_s(b\phi^2)\, ds$.

The existence and uniqueness in law of a solution to this martingale problem is well known [see Chapters 6 and 7 in Dawson (1993)], but as this and other references work with a larger class of test functions, we give a proof in the Appendix. See Theorem A.2. Let $P_{X_0}^{b,\sigma^2}$ denote the law of the solution on $\Omega_{X,C}$. We may, and shall at times, also consider this law as a probability on the space of cadlag paths $\Omega_{X,D}$.

We define a sequence of branching walk systems $\zeta_t^N$, $N = 1, 2, \dots$, using Brownian space–time scaling. Let $p(x, y) = p(y - x)$ be a fixed symmetric, finite range random walk kernel on $\mathbf{Z}^d$, with $p(0) = 0$ and with covariance matrix

$$\sum_{x \in \mathbf{Z}^d} x^i x^j p(x) = \delta_{i,j}\sigma^2,$$

where $\delta_{i,j} = 1$ if $i = j$, and $\delta_{i,j} = 0$ otherwise, is Kronecker's delta. Preparing for later generalizations we let $M_N = 1$ and define the rescaled lattice by

$$\mathbf{S_N} = \mathbf{Z}^d/(M_N\sqrt{N}) \equiv \{x/(M_N\sqrt{N}) : x \in \mathbf{Z}^d\}$$

and for $x \in \mathbf{S_N}$, define $p_N(x) = p(x\sqrt{N})$. Let $\zeta_t^N(x) = \zeta_{Nt}(x\sqrt{N})$ be the critical branching random walk system $\zeta_t^N : \mathbf{S_N} \to \mathbf{N}$ in which each particle dies at rate $\gamma N$ and produces a new particle at a given $y \in \mathbf{S_N}$ with rate $\gamma N p_N(y - x)$. Let $Y_t^N$ be the measure defined by putting an atom of size $1/N$ at each particle in $\zeta_t^N$. That is,

$$Y_t^N = \frac{1}{N} \sum_{x \in \mathbf{S_N}} \zeta_t^N(x)\delta_x. \tag{1.2}$$

For initial $\zeta_0^N$ with only finitely many particles, the law $P_N$ of $Y_\cdot^N$ is a probability measure on $\Omega_{X,D}$. We use $\Rightarrow$ to denote weak convergence on $\Omega_{X,D}$, or on the more mundane space $\mathbf{R}^d$. The following result is well known. See Theorem 4.6.2 of Dawson (1993), and note that his slightly different initial conditions are easy to modify to cover the present setting.

THEOREM 1.0. *Assume* $Y_0^N \Rightarrow Y_0 \in \mathscr{M}_F(\mathbf{R}^d)$ *as* $N \to \infty$. *Then* $P_N \Rightarrow P_{Y_0}^{2\gamma, \sigma^2\gamma}$ *as* $N \to \infty$.

We turn now to defining our rescaled voter models. We will create a somewhat general framework, so that we may easily consider a variety of different special cases. For $N = 1, 2, \ldots$, let $M_N \in \mathbf{N}$, and define $\mathbf{S_N}$ as above. We will use displacement distributions that depend on $N$. These, of course, will not be arbitrary but must satisfy various assumptions about their asymptotic behavior.

To formulate these assumptions, let $W_N = (W_N^1, \ldots, W_N^d) \in (\mathbf{Z}^d/M_N \setminus \{0\})$ have the displacement distribution $p$ on the unscaled lattice, and let $|x|$ denote the usual Euclidean norm of $x \in \mathbf{R}^d$. About the distribution of $W_N$ we will suppose that:

(H1)
- (a) $W_N$ and $-W_N$ have the same distribution.
- (b) There is a finite $\sigma^2 > 0$ such that $\lim_{N\to\infty} E(W_N^i W_N^j) = \delta_{ij}\sigma^2$.
- (c) The family $\{|W_N|^2, N \in \mathbf{N}\}$ is uniformly integrable.

The full strength of assumption (H1.a) is not needed. It would suffice to assume $EW_N = 0$, but symmetry simplifies many details in our proofs. We define the displacement distribution on the scaled lattice by

$$p_N(x) = P\left(\frac{W_N}{\sqrt{N}} = x\right), \qquad x \in \mathbf{S_N}, \tag{1.3}$$

and let $\xi_t^N(x) = \xi_{Nt}(x\sqrt{N})$ denote the rate-$N$ voter model on $\{0,1\}^{\mathbf{S_N}}$ with voting kernel $p_N(x, y) = p_N(y - x)$.

We treat $\xi_t^N$ as a measure by assigning mass $1/N'$ to each site of $\xi_t^N$ with value 1 and mass 0 to all other sites. Here the scaling for the particle mass satisfies $1 \le N' \le N$, and will depend on the particular model considered. Given a sequence $N'(N)$, we define the corresponding measure-valued process $X_t^N$ by

$$X_t^N = \frac{1}{N'} \sum_{x\in\mathbf{S_N}} \xi_t^N(x)\delta_x. \tag{1.4}$$

We make the following assumptions about the initial states $\xi_0^N$:

(H2)
- (a) $\sum_{x\in\mathbf{S_N}} \xi_0^N(x) < \infty$.
- (b) $X_0^N \to X_0$ in $\mathscr{M}_F(\mathbf{R}^d)$ as $N \to \infty$.

We note that a consequence of (H2) is that $\sup_N X_0^N(\mathbf{1}) < \infty$, a fact we will use frequently. Let $P_N$ denote the law of $X_\cdot^N$, and note that $P_N$ is a probability measure on $\Omega_{X,D}$.

The first special case of the above in which we will be interested is the following.

(M1) Long-range models. Let $M_N \to \infty$ as $N \to \infty$, and let $W_N$ be uniformly distributed on $(\mathbf{Z}^d/M_N) \cap I$, where $I = [-1, 1]^d \setminus \{0\}$.

Clearly, all the parts of (H1) are satisfied with $\sigma^2 = 1/3$.

The next result shows that for the long-range models defined above, $X_\cdot^N$ converges to the super-Brownian limit given in Theorem 1.0 with $\gamma = 1$ and $\sigma^2 = 1/3$ (the variance of the uniform distribution on $[-1, 1]$), provided the $p_N(\cdot)$ spread out rapidly enough.

THEOREM 1.1. *Assume* (H2) *holds, and let $P_N$ denote the law of $X_\cdot^N$ for the long-range models* (M1), *with $N' \equiv N$. If as $N \to \infty$,*

$$(1.5)\qquad \begin{aligned} M_N/\sqrt{N} &\to \infty \quad \text{in } d = 1,\\ M_N^2/(\log N) &\to \infty \quad \text{in } d = 2,\\ M_N &\to \infty \quad \text{in } d \geq 3. \end{aligned}$$

*then $P_N \Rightarrow P_{X_0}^{2,1/3}$ as $N \to \infty$.*

The above result for $d = 1$ should be compared to Theorem 2 of Mueller and Tribe (1995). They take $N' = N$ and $M_N = \sqrt{N}$ for $d = 1$ and obtain convergence of the approximate densities of $X_t^N$ to the solution of the stochastic partial differential equation (SPDE),

$$\frac{\partial u}{\partial t} = \frac{1}{6}\frac{\partial^2 u}{\partial x^2} + |2u(1-u)|^{1/2}\dot{W},$$

where $\dot{W}$ is a space–time white noise on $\mathbf{R}_+ \times \mathbf{R}$. The constants here and there do not match since our approximating voter models and limiting SPDE differ from theirs by a trivial scale factor. In the situation studied by Mueller and Tribe, the density of 0's and 1's are both nontrivial and are a continuous function of rescaled space.

It is well known that under $P_{Y_0}^{2\gamma,1/3}$, super-Brownian motion for $d = 1$ has a density which solves the above SPDE but with $\gamma$ in place of the $1 - u$ in front of the white noise [see Konno and Shiga (1988), Reimers (1989)]. Hence we can view the solution of the above SPDE as the density of a state dependent "super-Brownian motion" with branching parameter equal to $1 - u$, the local density of 0's for the limiting process. This agrees with our earlier description of the scaling limit of the voter models. By comparison, in the one-dimensional special case of Theorem 1.1, the fact that $M_N/\sqrt{N} \to \infty$ ensures that the local density of 0's is one in the limit, leading to branching with parameter $\gamma = 1$.

Durrett and Perkins (1999) have shown that, for $d \geq 2$, the scaling limit of super-critical contact process is super-Brownian motion with drift, provided the range of interaction $M_N \to \infty$ at the appropriate rate. Work of Derbez, Slade and Van der Hofstad (1998) on oriented percolation suggests that a corresponding limit theorem is valid for the critical contact process in $d > 4$, at least when the fixed kernel model is sufficiently spread out. We expect that their proof of this deep fact will be rather difficult and depend heavily on complicated lace expansion technology. Our computations for the convergence of rescaled fixed kernel voter models to super-Brownian motion at or above the critical dimension of 2 will be relatively simple, thanks to the existence of a dual process of coalescing random walks.

To study these voter models, we introduce our second set of examples:

(M2) Fixed Kernel Models. Let $M_N \equiv 1$, and let $p(x, y) = p(x - y)$ be an irreducible, symmetric, random walk kernel on $\mathbf{Z}^d$, such that $p(0) = 0$ and $\sum_{x \in \mathbf{Z}^d} x^i x^j p(x) = \delta_{ij}\sigma^2$. Define $W_N$ by $P(W_N = x) = p(x)$.

Clearly (H1) is satisfied in this case.

Our next result shows that for these fixed kernel models, one also obtains super-Brownian limits, but with a change in the limiting branching rate, and in $d = 2$ a different mass normalization $N'$. To specify the branching rate for the limit in $d \geq 3$, let $\gamma_e$ be the probability that a random walk with step distribution $p$, starting at the origin, never returns there.

THEOREM 1.2. *Assume* (H2) *holds, and let* $P_N$ *denote the law of* $X_\cdot^N$ *for the fixed kernel models* (M2), *with*

$$N' = \begin{cases} N/\log N, & \text{in } d = 2, \\ N, & \text{in } d \geq 3. \end{cases} \tag{1.6}$$

*Then* $P_N \Rightarrow P_{X_0}^{2\gamma, \sigma^2}$ *as* $N \to \infty$, *where*

$$\gamma = \begin{cases} 2\pi\sigma^2, & \text{in } d = 2, \\ \gamma_e, & \text{in } d \geq 3. \end{cases} \tag{1.7}$$

Consider the above result for $d \geq 3$ in the nearest neighbor case: $p(x) = 1/2d$ if $|x| = 1$. Although the overall density of 1's in space is still approximately zero, the local density of 0's, $V_t(x)$, defined in (1.1), at a site $x$ with $\xi_t(x) = 1$, is now the proportion of 0's among the nearest neighbors $y$ of $x$. This will be strictly less than 1 with significant probability, as a 1 at a neighbor $y$ may be due to a "birth" from $x$ which has not changed to 0 in the intervening time.

We will show that for $d \geq 3$ the mean proportion of neighboring 0's for an occupied site is $\gamma_e > 0$. The transience of simple random walk for $d \geq 3$ will allow us to conclude in addition that this nontrivial local density of 1's relies only on the contribution of "close cousins," and hence the resulting independence between sites at a positive macroscopic distance ensures there will be a mean-field simplification in the scaling limit. Note also that, unlike Theorem 1.0, this branching parameter does not affect the diffusion coefficient in Theorem 1.2. This is because for the $\{0, 1\}$-valued voter model we can also exchange neighboring 1's (it won't change a thing) and so the effective diffusion rate is given by the variance of the voter kernel. To see this more clearly, the reader may want to look at the derivation of (2.4) in the proof of Theorem 2.1(i) below.

If we take $N' = N$ in the fixed kernel case in $d = 2$, then the recurrence of simple random walk ensures that the local density of 0's about a site in state 1 approaches zero, and the resulting limit will be super-Brownian motion with no branching, that is, deterministic heat flow. Presutti and Spohn (1983) stated this result at the end of their Section 2, but did not give the details of the proof. Cox and Durrett (1995) (see their Theorem 2) give a complete proof for the special initial condition of a half plane of 1's, which can be easily adapted to more general initial macroscopically smooth initial conditions. From the last two results, we see that to get a random limit, we must compensate for the low density of 0's in $d = 2$, by speeding up the branching rate $N$, or, equivalently, reduce the inverse particle mass $N'$, which has been done in Theorem 1.2.

Theorem 1.2 has been applied in Bramson, Cox and Le Gall (1999) to resolve a conjecture raised in Bramson and Griffeath (1980). In that paper, the asymptotic behavior of $P(\xi_t \neq \emptyset)$ as $t \to \infty$ was obtained, where $\xi_t$ is the nearest neighbor voter model on $\mathbb{Z}^d$ with rate one started from a single one at the origin. Bramson and Griffeath asked whether or not $\xi_t$, conditioned on the event $\{\xi_t \neq \emptyset\}$, obeyed an asymptotic shape theorem. Using Theorem 1.2, Bramson, Cox and Le Gall answer this by showing that the law of $\xi_t/\sqrt{t}$, conditioned on the event $\{\xi_t \neq \emptyset\}$, converges as $t \to \infty$ to the law of the support of super-Brownian motion at time $t = 1$ under its canonical measure [see Section 3.4 of Dawson (1992)].

Finally, we consider the case $d = 1$ which is not covered by Theorem 1.2. If we restrict our attention to the nearest neighbor case, then taking $N' = N^{1/2}$, and using the reasoning that led to Theorem 2 of Cox and Griffeath (1986), we see that $X_t^N$ defined above converges to a measure valued process where the density at any positive time is 1 or 0 on alternating intervals of random length, with the end points of these intervals undergoing annihilating Brownian motions.

Having considered fixed range and long-range voter models, it is natural to ask if there are any results to be found in between. In $d \geq 3$, Theorems 1.1 and 1.2 cover all the possibilities. The Mueller and Tribe (1995) result shows that at the borderline of the $d = 1$ condition in Theorem 1.1, we get a more interesting limit. Extrapolating wildly suggests that in $d = 2$ we should take $M_N = O(\sqrt{\log N})$ to get a more interesting limit. The next result does this and yields a result which "interpolates" between Theorems 1.1 and 1.2.

THEOREM 1.3. *Assume* (H2) *holds, and let* $P_N$ *denote the law of* $X_\cdot^N$ *for the* $d = 2$ *long range models* (M1), *where as* $N \to \infty$, $M_N \to \infty$ *and*

$$\frac{M_N^2}{\log N} \to \rho \in [0, \infty), \tag{1.8}$$

$$N' = \begin{cases} N, & \rho > 0, \\ NM_N^2/\log N, & \rho = 0. \end{cases} \tag{1.9}$$

*Then* $P_N \Rightarrow P^{2\gamma,1/3}_{X_0}$ *as* $N \to \infty$, *where*

$$\gamma = \begin{cases} 1/[1+(3/2\pi\rho)], & \rho > 0, \\ 2\pi/3, & \rho = 0. \end{cases} \tag{1.10}$$

We note that setting $\rho = \infty$ in (1.10) gives $\gamma = 1$, which is consistent with Theorem 1.1. For any $\rho < \infty$ we have $0 < \gamma < 1$. To explain this, note that the difference of two independent random walks with variance $v$ will visit a cube about $\log N/2\pi v$ times up to time $N$, so if the number of lattice points in the cube $M_N^2 = O(\log N)$ there is positive probability of the difference hitting 0 (and positive probability it does not).

As $\rho \to 0$ the probability the two random walks hit approaches 1, so the local density of 0's approaches 0 and we have to reduce the initial mass to $NM_N^2/\log N$. To see the result of this, note that if we used $N' = M_N^2/\log N$ in the case $\rho > 0$, that would change the $\rho > 0$ formula for $\gamma$ to

$$\frac{1/\rho}{1+(3/2\pi\rho)} \to \frac{2\pi}{3} \quad \text{as } \rho \to 0.$$

Finally, we should consider $d = 1$. Theorem 1.1 applies if $M_N/\sqrt{N} \to \infty$, while Mueller and Tribe handle the case $M_N/\sqrt{N} \to \rho$. We leave it to the interested reader to show that if $M_N/\sqrt{N} \to 0$ then we get the same result as for the nearest neighbor case we discussed after Theorem 1.2.

To minimize the total number of words needed to write down our proofs and to better expose the reasoning behind our arguments, Theorems 1.1–1.3 will be derived as special cases of more general results. In addition to (H1) and (H2), which concern the dispersal kernel and the initial condition, we will need the following assumptions about the mass renormalization:

(H3)
(a) $1 \le N' \le N$,
(b) $N' \equiv N$ or $N'/N \to 0$,
(c) $\lim_{N\to\infty} N^{5/7}/N' = 0$.

The first two conditions are natural but the third is somewhat strange. The 5/7's here is dictated by the proof of (2.13) in Lemma 2.4. In all of the examples we consider, we will either have $N' \equiv N$ or $N/N' = \log N$, so we could get away with any power less than 1.

Let $B_t^{x,N} \in \mathbf{S_N}$ be independent rate $N$ random walks with $B_0^{x,N} = x$ and step distribution $p_N(\cdot)$. To prove our results, we will need assumptions on the behavior of $B_t^{x,N}$ and of $B_t^{*,N}$, the rate $N$ random walk with initial distribution $p_N(\cdot)$. Let

$$\tau^{*,N} = \inf\{t \ge 0 : B_t^{*,N} = 0\} \le \infty.$$

Our new conditions are as follows:

$$\lim_{N\to\infty} \frac{N}{N'} P(\tau^{*,N} > t) = \gamma. \tag{K1}$$

(K2) There is a sequence $\varepsilon_N \to 0$ with $0 < \varepsilon_N < 1/2$, so that for all $t > 0$ and $\delta > 0$,

$$\text{(a)} \qquad \lim_{N\to\infty} \frac{N}{N'} P\Big(\big|B^{0,N}_{t\varepsilon_N}\big| > \delta\Big) = 0,$$

$$\text{(b)} \qquad \lim_{N\to\infty} \frac{N}{N'} P(\varepsilon_N t < \tau^{*,N} \le t) = 0.$$

$$\text{(K3)} \qquad \lim_{N\to\infty} \sup_{x\in \mathbf{S_N}} N' P(B^{0,N}_t = x) = 0 \quad \text{for all } t > 0.$$

$$\text{(K4)} \qquad \sup_N \frac{N}{N'} P\Big(\tau^{*,N} > (N'/N)^2\Big) < \infty.$$

Condition (K1) says that the ratio $N/N'$ was chosen correctly and identifies the value of the branching rate $\gamma$. (K3) says that $N'$ is not too large or the lattice spacing $M_N^{-1} M^{-1/2}$ is not too large. This condition fails, for example, in $d = 1$ if $N' = N$ and $M_N/\sqrt{N} \not\to \infty$, as is the case in Theorem 2 of Mueller and Tribe (1995) described above and for which the weak limit is not super-Brownian motion. To get a sense for (K2) and (K4), first consider the case $N' = N$. In this situation, (K4) is trivial while the condition in (K2) involves a competition between wanting $\varepsilon_N \to 0$ quickly or slowly. In the case $N' = N$ this is typically easy to satisfy but when $N'/N \to 0$, (K4) becomes nontrivial and the competition in (K2) becomes more difficult to accommodate.

The conditions (K1)–(K4) are designed to allow us to prove the following umbrella result.

THEOREM 1.4. *Assume* (H1)–(H3) *and* (K1)–(K4). *Then* $P_N \Rightarrow P^{2\gamma,\sigma^2}_{X_0}$.

The proof of Theorem 1.4 begins in Section 2 with a "stochastic integral" construction of the voter model which was used by Mueller and Tribe (1995) and by Kurtz and Protter (1996). In Section 2 we also identify certain related martingales and their square functions. In Section 3 we introduce a set of intermediate conditions (I1)–(I3) and prove the following theorem.

THEOREM 1.5. *Assume* (H1)–(H3) *and* (I1)–(I3). *Then* $P_N \Rightarrow P^{2\gamma,\sigma^2}_{X_0}$.

In Section 4, we show that in the presence of (H1)–(H3), (K1)–(K4) imply (I1)–(I3). In Section 5 we verify in Theorem 5.1 that (K1)–(K4) hold for the voter models considered in Theorems 1.1, 1.2 and 1.3, thus completing their proofs. In our first version of this paper, we only considered dimensions $d \ge 2$. We thank Martin Barlow for suggesting that we extend our results to the one-dimensional setting, for this has led to an illuminating comparison between $d \ge 3$, $d = 2$ and $d = 1$. That is, Theorems 1.1–1.3 imply that the limit is always super-Brownian motion in dimensions $d \ge 2$. In contrast, in $d = 1$ we get super-Brownian motion, a SPDE, or a $\{0, 1\}$-valued density depending on the size of the range.

**2. Rescaled voter models and martingale problems.** In this section, we give a careful construction of our rescaled voter models $\xi_t^N$ and describe associated measure-valued martingale problems solved by their empirical measures $X_t^N$. The main result here is Theorem 2.1 below.

The rescaled voter model $\xi_t^N \in \{0,1\}^{\mathbf{S_N}}$, described in the introduction, evolves as follows. Each site $x \in \mathbf{S_N}$, at rate $N$, selects a site $y \in \mathbf{S_N}$ with probability $p_N(y-x)$, and adopts the "opinion" of the chosen site. Transitions at different sites and times are independent of one another. In order to make effective use of martingale methods, we follow the approach of Mueller and Tribe (1995) [see also Kurtz and Protter (1996)] and give a more formal construction of $\xi_t^N$ using $\xi_0^N$ and a family of independent Poisson processes $\{\Lambda_t^N(x,y)\colon x,y \in \mathbf{S_N}\}$, with rate $N\,p_N(y-x)$.

To be very formal, these Poisson processes are defined on a complete probability space and we let $\mathscr{F}_t$ denote the canonical right-continuous filtration generated by these processes up to time $t$ and the $P$-null sets. To explain the phrase "rate $N\,p_N(y-x)$" we note that the compensated processes,

$$\hat{\Lambda}_t^N(x,y) = \Lambda_t^N(x,y) - N\,p_N(y-x)t,$$

are $(\mathscr{F}_t)$-martingales. Our rescaled voter model is defined by the stochastic integral equation,

$$\xi_t^N(x) = \xi_0^N(x) + \sum_y \int_0^t \left[\xi_{s-}^N(y) - \xi_{s-}^N(x)\right] d\Lambda_s^N(x,y), \tag{2.1}$$

$$x \in \mathbf{S_N},\ t \ge 0,$$

where the sum is over $y \in \mathbf{S_N}$. A solution $\xi^N$ is a cadlag $\{0,1\}^{\mathbf{S_N}}$-valued process for which

$$\sum_y \int_0^t \left|\xi_{s-}^N(y) - \xi_{s-}^N(x)\right| d\Lambda_s^N(x,y) < \infty \quad \text{for all } t>0,\ x \text{ a.s.}$$

and (2.1) holds a.s. To see that this gives the voter model, note that if $\xi_{s-}^N(y) = \xi_{s-}^N(x)$ then nothing happens, then consider the cases $\xi_{s-}^N(y) = 1$, $\xi_{s-}^N(x) = 0$ and $\xi_{s-}^N(y) = 0$, $\xi_{s-}^N(x) = 1$.

Equation (2.1) captures our earlier more intuitive description, but, as it involves infinitely many processes $\xi_\cdot^N(x)$, $x \in \mathbf{S_N}$, we will give here an elementary proof of the existence of a unique solution. In doing this, we will omit the dependence on $N$ from our notation [writing $\xi_t$ for $\xi_t^N$, $\hat{\Lambda}_t(x,y)$ for $\hat{\Lambda}_t^N(x,y)$, etc.]. We will often do this in what follows except where confusion might arise. A more general version of the following lemma may be found in Chapter 9 of Kurtz and Protter (1996).

LEMMA 2.1. *With probability* 1, *equation* (2.1) *uniquely defines* $\xi_t$ *for initial states* $\xi_0$ *with finitely many sites having opinion* 1.

PROOF. Let $T_0 = 0$ and for $n \ge 1$, let $T_n$ be the time of the $n$th jump of $\sum_{x,y} \xi_t(x)(\Lambda_t(x,y) + \Lambda_t(y,x))$. In words, $T_1$ is the time until the first jump of some $\Lambda_t(x, y)$ or $\Lambda_t(y, x)$ with $\xi_t(x) = 1$. Since $K_N \equiv \sum_x \xi_0(x) < \infty$, and $\sum_z p(z) = 1$, $T_1$ is an exponential random variable with mean $\mu_1 = (2NK_N)^{-1}$. Clearly, (2.1) uniquely defines $\xi_t$ for $t \in [0, T_1)$, since $\xi_t$ is constant for such $t$. Furthermore, it is easy to see by induction that (2.1) has a unique solution on $[0, T_\infty)$ where $T_\infty = \lim_{n\to\infty} T_n$, so it suffices now to show that $T_\infty = \infty$ a.s.

To get an upper bound on the number of 1's at time $t$, we set

$$Y_t = K_N + j \quad \text{for } T_j \le t < T_{j+1}.$$

The time difference $T_{j+1} - T_j$ is bounded below by an exponential random variable with mean $(2N(K_N + j))^{-1}$. Thus if $Z_t$ defines a pure birth process in which each particle gives birth at rate $2N$, we can define $Y_t$ and $Z_t$ on the same space so that $Y_t \le Z_t$. Well-known results for the pure birth process imply $Z_t < \infty$ a.s. for all $t$, and hence $T_\infty = \infty$ a.s. This can be proved for instance by using a trivial comparison and then computing the mean of the branching process

$$E\Big(\sup_{t\le T} \sum_x \xi_t(x)\Big) \le E(Z_T) = K_N \exp(2NT) < \infty. \tag{2.2}$$

This estimate will be useful later on. □

Recall that for bounded $\phi: \mathbf{S_N} \to \mathbf{R}$, we have defined in (1.4),

$$X_t^N(\phi) = \frac{1}{N'} \sum_x \phi(x)\xi_t(x).$$

For bounded, measurable $\psi: [0, T] \times \mathbf{S_N} \to \mathbf{R}$, let

$$M_t^N(\psi) = \frac{1}{N'} \sum_x \sum_y \int_0^t \psi_s(x)(\xi_{s-}(y) - \xi_{s-}(x))\, d\hat{\Lambda}_s(x, y), \qquad 0 \le t \le T. \tag{2.3}$$

Note that (2.2) shows that this sum is finite for all $t \in [0, T]$ a.s.

We need a little more notation before stating Theorem 2.1. Recall that, for $x \in \mathbf{S_N}$, $B_t^{x,N}$ denotes the continuous time random walk on $\mathbf{S_N}$ which starts at $x$, jumps at rate $N$ and has step distribution $p_N$. Let $P_t^N f(x) = Ef(B_t^{x,N})$ be its semigroup and $\mathscr{A}_N$ its generator,

$$\mathscr{A}_N \phi(x) = N \sum_y p_N(y - x)\big(\phi(y) - \phi(x)\big).$$

The walks $B_t^{x,N}$, $x \in \mathbf{S_N}$ are independent. Let $Z_t^{x,N} = B_{t/N}^{x,N}\sqrt{N}$ be the corresponding "unscaled" walks. For $\phi: [0, T] \times \mathbf{S_N} \to \mathbf{R}$, we let $\dot{\phi}(s, x) = (\partial\phi/\partial s)(s, x)$, and write $\phi_t(x)$ for $\phi(t, x)$ and $\dot{\phi}_t(x)$ for $\dot{\phi}(t, x)$. Let $C_b^{m,n}([0, T] \times \mathbf{R}^d)$ be the space of bounded continuous functions on $[0, T] \times \mathbf{R}^d$ whose derivatives of order less than $m + 1$ in the first variable and partial derivatives of order less than $n + 1$ in the second variable are also bounded and continuous. [We allow $T = \infty$, in which case $\phi$ is taken to have domain

$[0,\infty)\times\mathbf{S_N}$.] Let $\mathbf{1}$ denote the function on $\mathbf{S_N}$ which is identically one. Finally, we introduce two notions of the density of vacant sites near $x$ at time $t$,

$$V_N(t,x)=\sum_y p_N(y-x)1\{\xi_t(y)=0\},\qquad V'_N(t,x)=\frac{N}{N'}V_N(t,x)$$

and define $V_{N,t}(x)=V_N(t,\,x)$, $V'_{N,t}(x)=V'_N(t,x)$.

THEOREM 2.2. *Let* $\phi\in C_b^{1,3}([0,\infty)\times\mathbf{R}^d)$. *Then*

(i)
$$X_t^N(\phi)=X_0^N(\phi)+\int_0^t X_s^N(\dot\phi_s+\mathscr{A}_N\phi_s)\,ds+M_t^N(\phi),$$

*where*

(ii) $M_t^N(\phi)$ *[defined in (2.3)] is a cadlag, square-integrable* $(\mathscr{F}_t)$*-martingale, with predictable square function*

$$\langle M^N(\phi)\rangle_t=\int_0^t[2X_s^N(\phi_s^2V'_{N,s})+\varepsilon_s^N(\phi)]\,ds,$$

*where* $\varepsilon_s^N(\phi)$ *satisfies*

$$E\Big(\sup_{s\le T}|\varepsilon_s^N(\phi)|^2\Big)\to 0\quad \textit{as } N\to\infty \textit{ for any } T>0.$$

(iii) *For any* $T>0$,

$$E\int_0^T X_s^N(|\mathscr{A}_N\phi_s-\sigma^2\Delta\phi_s/2|)\,ds\to 0\quad \textit{as } N\to\infty.$$

OUTLINE OF PROOF. Conclusion (i) is (2.5) below. To prove (ii), we will begin in Lemma 2.2 by deriving a formula for $\langle M^N(\psi)\rangle_t$ for a large class of test functions $\psi$, and then analyze the behavior of $M^N$ in Lemma 2.4. Finally, the convergence in (iii) of the rescaled random walk generators to the Laplacian is proved in Lemma 2.5.

PROOF OF (i). Let $\phi\colon[0,T]\times\mathbf{S_N}\to\mathbf{R}$ such that both $\phi$ and $\dot\phi(s,x)$ are in $C_b([0,T]\times\mathbf{S_N})$. It follows from the integration by parts formula of Riemann–Stieltjes integration theory and (2.1) that, for $t\in[0,T]$,

$$\begin{aligned}\xi_t(x)\phi_t(x)&=\xi_0(x)\phi_0(x)+\int_0^t\xi_s(x)\dot\phi_s(x)\,ds\\&\quad+\sum_y\int_0^t\phi_s(x)(\xi_{s-}(y)-\xi_{s-}(x))\,d\hat\Lambda_s(x,y)\\&\quad+\sum_y\int_0^t\phi_s(x)(\xi_{s-}(y)-\xi_{s-}(x))\,Np_N(y-x)\,ds.\end{aligned}$$

Note that the left limits are important in the second integral on the right but not in the third where the integrator is continuous. Multiplying by $1/N'$, summing over $x$ and rearranging gives

$$X_t^N(\phi_t) = X_0^N(\phi_0) + M_t^N(\phi)$$
$$(2.4) \qquad + \int_0^t \left[ X_s^N(\dot{\phi}_s) + \frac{N}{N'}\left( \sum_x \sum_y (\xi_s(y) - \xi_s(x)) p_N(y-x) \right) \phi_s(x) \right] ds,$$

where, as in (2.3),

$$M_t^N(\phi) = \frac{1}{N'} \sum_x \sum_y \int_0^t \phi_s(x)(\xi_{s-}(y) - \xi_{s-}(x))\, d\hat{\Lambda}_s(x, y),$$

and absolute convergence of this series and all the sums in (2.4) is clear from (2.2). By summation by parts and the symmetry assumption (H1)(a),

$$\frac{N}{N'} \sum_x \sum_y (\xi_s(y) - \xi_s(x)) \phi_s(x) p_N(y-x)$$
$$= \frac{1}{N'} \sum_x \xi_s(x) \left( N \sum_y (\phi_s(y) - \phi_s(x)) \right) p_N(y-x) = X_s^N(\mathscr{A}_N \phi_s).$$

Therefore, (2.4) becomes

$$(2.5) \quad X_t^N(\phi_t) = X_0^N(\phi_0) + \int_0^t X_s^N(\dot{\phi}_s + \mathscr{A}_N \phi_s)\, ds + M_t^N(\phi), \qquad t \in [0, T].$$

PROOF OF (ii). Having obtained this representation for $X_t^N$, we must now show that $M_t^N(\phi)$ is a martingale, and obtain the required form for its predictable square function. The first step in doing so is the following lemma.

LEMMA 2.3. *Fix $T < \infty$, and let $\psi: [0, T] \times \mathbf{S_N} \to \mathbf{R}$ be bounded and measurable. Then $M_t^N(\psi)$ is a cadlag, square-integrable $(\mathscr{F}_t)$-martingale, with predictable square function*

$$(2.6) \quad \langle M^N(\psi) \rangle_t = \frac{N}{N'^2} \int_0^t \sum_x \sum_y \psi_s(x)^2 (\xi_s(y) - \xi_s(x))^2 p_N(y-x)\, ds, \; t \in [0, T].$$

*In particular*,

$$(2.7) \qquad \langle M^N(\mathbf{1}) \rangle_t = 2 \int_0^t X_s^N(V'_{N,s})\, ds, \qquad t \geq 0.$$

PROOF. Equation (2.7) follows from (2.6) and the relevant definitions. To prove (2.6) we will truncate to a finite sum, apply standard results for stochastic integrals with respect to Poisson processes and then pass to the limit. For $k = 1, 2, \ldots$, let

$$M_t^{N,k}(\psi) = \sum_{x: |x| \leq k} \sum_{y: |y| \leq k} \int_0^t \frac{\psi_s(x)}{N'} (\xi_{s-}(y) - \xi_{s-}(x))\, d\hat{\Lambda}_s^N(x, y).$$

Clearly, $M_t^{N,k}(\psi)$ is a cadlag, square-integrable $(\mathscr{F}_t)$-martingale with predictable square function

$$\langle M^{N,k}(\psi)\rangle_t = \int_0^t \sum_{x:|x|\le k} \sum_{y:|y|\le k} \frac{\psi_s(x)^2}{N'^2} (\xi_s(y) - \xi_s(x))^2 \, N p_N(y-x)\, ds.$$

Since the difference $\xi_s(y) - \xi_s(x)$ can only be nonzero when either $\xi_s(x) = 1$ or $\xi_s(y) = 1$, it is easy to see that there is a finite constant $C_N$ so that

$$E\langle M^{N,k}(\psi)\rangle_t \le C_N \|\psi\|_\infty^2 \int_0^t E(X_s^N(\mathbf{1}))\, ds < \infty$$

by (2.2). Since $k \to \langle M^{N,k}(\psi)\rangle_t$ is increasing, the last estimate and dominated convergence imply

$$\sup_{k,\, j\ge K} E\langle M^{N,k}(\psi) - M^{N,j}(\psi)\rangle_T \to 0 \quad \text{as } K \to \infty,$$

and so by the $L^2$ maximal inequality for martingales,

$$\sup_{k,\, j\ge K} E\Big(\sup_{t\le T} |M_t^{N,k}(\psi) - M_t^{N,j}(\psi)|^2\Big) \to 0 \quad \text{as } K \to \infty.$$

Letting $j \to \infty$ and using Fatou's lemma, converts this into

$$E\Big(\sup_{t\le T} |M_t^{N,k}(\psi) - M_t^{N}(\psi)|^2\Big) \to 0 \quad \text{as } k \to \infty.$$

This shows that $M_t^N(\psi)$ is a cadlag (right-continuous with left limits), $L^2$ $(\mathscr{F}_t)$-martingale, and it follows easily that $\langle M^N(\psi)\rangle_t = \lim_{k\to\infty}\langle M^{N,k}(\psi)\rangle_t$ which gives the desired formula. □

The next result is a technical interlude needed to compute the mean of $X_t^N(\psi)$ and get bounds on its moments.

LEMMA 2.4. (a) *If* $\psi\colon \mathbf{S_N} \to \mathbf{R}$ *is bounded, then* $EX_t^N(\psi) = X_0^N(P_t^N\psi)$. *In particular*,

$$EX_t^N(\mathbf{1}) = X_0^N(\mathbf{1}).$$

(b) *For any* $p > 1$ *and* $T > 0$, *there is a finite constant* $c_{p,T}$ *such that*

$$E\Big(\sup_{t\le T} X_t^N(\mathbf{1})^p\Big) \le c_{p,T}(N/N')^{p-1/2}(X_0^N(\mathbf{1})^p + 1).$$

PROOF. (a) Set $\phi_s(x) = P_{t-s}^N\psi(x)$ for $0 \le s \le t$ and note that $\dot\phi_s + \mathscr{A}_N\phi_s = 0$ by the backwards equation for continuous time Markov chains. Now take expectation in (2.5) at $t = T$ and use the fact that $M_t^N(\phi)$ is a martingale.

(b) Lemma 2.2 and (2.5) show that $X_t^N(\mathbf{1}) = X_0^N(\mathbf{1}) + M_t^N(\mathbf{1})$ is an $L^2$ martingale, such that

$$\text{(2.8)} \qquad \langle X^N(\mathbf{1})\rangle_t = \frac{2N}{N'}\int_0^t \frac{1}{N'}\sum_x \xi_s(x) V_N(s,x)\, ds \le \frac{2N}{N'}\int_0^t X_s^N(\mathbf{1})\, ds,$$

where the second equality follows from $V_N(s,x) \le 1$. For $p > 1$, a predictable square function inequality of Burkholder (1973) (see Theorem 21.1) shows that, for a finite constant $b_p$,

$$E\Big(\sup_{t\le T} X_t^N(\mathbf{1})^p\Big) \le b_p\Big[(X_0^N(\mathbf{1}))^p + E\langle X^N(\mathbf{1})\rangle_t^{p/2} + E\Big(\sup_{s\le T} |\Delta X_s^N(\mathbf{1})|^p\Big)\Big]. \tag{2.9}$$

Since the largest discontinuity in the process $X_s^N(\mathbf{1})$ is at most $1/N' \le 1$ and the integral representing $\langle X^N(1)\rangle_t$ in (2.8) can be trivially bounded by the supremum, we have

$$E\Big(\sup_{t\le T} X_t^N(\mathbf{1})^p\Big) \le b_p\Big[X_0^N(\mathbf{1})^p + \Big(\frac{2NT}{N'}\Big)^{p/2} E\Big(\sup_{t\le T} X_t^N(\mathbf{1})^{p/2}\Big) + 1\Big]. \tag{2.10}$$

Inequality (2.10) will reduce the desired result for $p$ to the result for $p/2$. To get the induction going, we prove the conclusion for $1 < p \le 2$. Setting $p = 2$ in (2.9) and using (2.8) with conclusion (a) of this lemma we have

$$E\Big(\sup_{t\le T} X_t^N(\mathbf{1})^2\Big) \le b_2\Big[X_0^N(\mathbf{1})^2 + (2TN/N') + 1\Big] \le b_2(2T+1)\frac{N}{N'}\Big(X_0^N(\mathbf{1})^2 + 1\Big)$$

since $N/N' \ge 1$. For $1 < p \le 2$, it follows from Jensen's inequality and the fact that $p/2 \le p - 1/2$, that

$$E\Big(\sup_{t\le T} X_t^N(\mathbf{1})^p\Big) \le c_{p,T}\Big(\frac{N}{N'}\Big)^{p-1/2}(X_0^N(\mathbf{1})^p + 1)$$

for a finite constant $c_{p,T}$.

We use induction to finish the proof. Suppose that (b) has been proved for $2^{n-1} < p \le 2^n$ with $n \ge 1$, and we let $2^n < p \le 2^{n+1}$. Using (2.10) and the induction hypothesis,

$$E\Big(\sup_{t\le T} X_t^N(\mathbf{1})^p\Big) \le b_p\Big[X_0^N(\mathbf{1})^p + \Big(\frac{2NT}{N'}\Big)^{p/2} c_{p/2,T}\Big(\frac{N}{N'}\Big)^{p/2-1/2}(X_0^N(\mathbf{1})^{p/2} + 1) + 1\Big],$$

which gives the desired result. □

REMARK. The estimate in (b) above is poor if $N' \ll N$. For this situation, in the models we consider, better bounds are derived in Section 4 for $p = 2$ and 3.

To prepare for taking the limit as $N \to \infty$, we will now rewrite the expression for $\langle M^N(\psi)\rangle_t$ given in Lemma 2.2.

LEMMA 2.5. *Let* $\psi: [0, T] \times \mathbf{S_N} \to \mathbf{R}$ *be bounded and measurable, such that* $\psi^2$ *is Lipschitz continuous on compact time sets; that is,*

$$\left|\psi^2(s, x) - \psi^2(s, y)\right| \le C_\psi |x - y|, \qquad s \in [0, T], \qquad x, y \in \mathbf{S_N},$$

*for a finite constant* $C_\psi$. *In this case,*

$$\langle M^N(\psi)\rangle_t = \int_0^t \left[2X_s^N(\psi_s^2 V'_{N,s}) + \varepsilon_s^N(\psi)\right] ds, \tag{2.11}$$

*where, for another finite constant* $C_\psi$,

$$\left|\varepsilon_s^N(\psi)\right| \le C_\psi \frac{\sqrt{N}}{N'} X_s^N(\mathbf{1}) \tag{2.12}$$

*and*

$$E\left(\sup_{s \le T} |\varepsilon_s^N(\psi)|^2\right) \to 0 \quad \textit{as } N \to \infty. \tag{2.13}$$

PROOF. First we claim that for bounded $\phi: \mathbf{S_N} \to \mathbf{R}$ we have

$$\begin{aligned} \frac{1}{N'} \sum_x \sum_y \phi(x)(\xi_s(y) - \xi_s(x))^2 p_N(y - x) \\ = 2X_s^N(\phi V_{N,s}) \\ + \frac{1}{N'} \sum_x \sum_y (\phi(x) - \phi(y)) p_N(y - x) \xi_s(y)(1 - \xi_s(x)). \end{aligned} \tag{2.14}$$

To see this, begin with the identity

$$(\xi_s(y) - \xi_s(y))^2 = \xi_s(x) 1(\xi_s(y) = 0) + \xi_s(y) 1(\xi_s(x) = 0).$$

Using this, the left-hand side of (2.14) can be written as

$$\begin{aligned} \frac{1}{N'} \sum_x \phi(x) \xi_s(x) V_{N,s}(x) + \frac{1}{N'} \sum_y \phi(y) \xi_s(y) V_{N,s}(y) \\ + \frac{1}{N'} \sum_x \sum_y (\phi(x) - \phi(y)) p_N(y - x) \xi_s(y) 1(\xi_s(x) = 0), \end{aligned}$$

which equals the right side of (2.14).

Consulting (2.6), then letting $\phi = \psi_s^2$ in (2.14) and multiplying by $N/N'$, shows that (2.11) holds, with

$$\varepsilon_s^N(\psi) = \frac{N}{N'^2} \sum_x \sum_y (\psi_s(x)^2 - \psi_s(y)^2) p_N(y - x) \xi_s(y) 1(\xi_s(x) = 0).$$

By our assumption on $\psi$, and the scaling of $p_N$ in (1.3),

$$\left|\varepsilon_s^N(\psi)\right| \le \frac{N}{N'^2} \sum_y \xi_s(y) \sum_x C_\psi |x - y| p_N(y - x) \le C_\psi \frac{N}{N'} X_s^N(\mathbf{1}) \frac{1}{N^{1/2}} E(|W_N|).$$

By (H1.c), $E|W_N|$ is bounded, and hence (2.12) must hold for some finite $C_\psi$.

Finally, to prove (2.13) we note that from (2.12), and Lemma 2.3(b),

$$E\sup_{s\le T}\left|\varepsilon_s^N(\psi)\right|^2 \le C_\psi \frac{N}{(N')^2} E\sup_{s\le T} X_s^N(\mathbf{1})^2 \le C_\psi \frac{N^{5/2}}{(N')^{7/2}}(X_0^N(\mathbf{1})^2+1) \to 0$$

by our basic assumptions (H2) and (H3)(c). □

PROOF OF (iii). The next result shows that the rescaled random walk generators $\mathscr{A}_N$ converge in an appropriate sense to the generator of Brownian motion. As the reader can probably guess, this is a straightforward application of Taylor's theorem with remainder.

LEMMA 2.6. *For* $\phi \in C_b^{1,3}([0,T]\times \mathbf{R}^d)$,

$$\lim_{N\to\infty}\sup_{s\le T}\|\mathscr{A}_N\phi_s - \sigma^2\Delta\phi_s/2\|_\infty = 0.$$

*Moreover, for every* $R<\infty$, *the rate of convergence above is uniform over*

$$\Big\{\phi\in C_b^{1,3}([0,T]\times\mathbf{R}^d)\colon \sup_{s,i,j,k}(\|(\phi_s)_{ij}\|_\infty + \|(\phi_s)_{ijk}\|_\infty)\le R\Big\},$$

where the subscripts $i, j, \dots$ indicate partial derivatives with respect to the spatial variable.

PROOF. Taylor's theorem shows there is a $Y_N(\omega)$ in the line segment from $x$ to $x+W_N(\omega)/\sqrt{N}$ such that

$$N\big(\phi_s(x+W_N/\sqrt{N})-\phi_s(x)\big) = \nabla\phi_s(x)(W_N/\sqrt{N}) + \tfrac12\sum_{ij}(\phi_s)_{ij}(Y_N)W_N^iW_N^j.$$

Therefore, using (H1), we have

$$\begin{aligned}\big|\mathscr{A}_N\phi_s(x)-\sigma^2\Delta\phi_s(x)/2\big| &= \big|N(E\phi_s(x+W_N/\sqrt{N})-\phi_s(x))-\sigma^2\Delta\phi_s(x)/2\big|\\ &\le \tfrac12\sum_{i,j}\Big|E\Big(\{(\phi_s)_{ij}(Y_N)-(\phi_s)_{ij}(x)\}W_N^iW_N^j\Big)\Big|\\ &\quad+\tfrac12\sum_{i,j}\Big|(\phi_s)_{ij}(x)(E(W_N^iW_N^j)-\delta_{ij}\sigma^2)\Big|.\end{aligned}$$

Since $(\phi_s)_{ij}$ is bounded and Lipschitz continuous, the above is not larger than

$$C\sum_{ij}E\bigg(\bigg(\frac{|W_N|}{\sqrt{N}}\wedge 1\bigg)|W_N^iW_N^j|\bigg) + C\max_{1\le i,\,j\le d}\big|E(W_N^iW_N^j)-\delta_{ij}\sigma^2\big| \equiv \eta_N$$

for a finite constant $C$. By (H1)(b) and (H1)(c), $\eta_N\to 0$ uniformly as required. □

This, together with Lemma 2.3(a) and (H2), completes the proof of (iii) in Theorem 2.1, which means that the mission of this section has been accomplished.

**3. Weak convergence to super-Brownian motion.** We now introduce some intermediate hypotheses, (I1)–(I3). In this section, we will show that (H1)–(H3) and (I1)–(I3) imply the conclusion of Theorem 1.4. In Section 4 we will confront the problem of inferring (I1)–(I3) from (K1)–(K4). The three new conditions are:

(I1) There is a finite $\gamma > 0$ such that, for all $\phi \in C_0^\infty(\mathbf{R}^d)$ and $T > 0$, as $N \to \infty$,

$$E\left[\left(\int_0^T X_s^N(\{V'_{N,s} - \gamma\}\phi^2)\,ds\right)^2\right] \to 0.$$

(I2) For all $T > 0$ there exists a finite $C_T$ such that $\lim_{T\downarrow 0} C_T = 0$ and for all $N$,

$$\int_0^T E^{\xi_0^N}[X_s^N(V'_{N,s})]\,ds \le C_T X_0^N(\mathbf{1}).$$

(I3) There is a $\theta \in (0, 1]$ and a finite $C(\varepsilon, T, K)$ such that, for all $N \in \mathbf{N}$, all cutoffs $0 < \varepsilon$, $K < \infty$, and all pairs of times $\varepsilon \le s \le t \le T$, we have

$$\sup\left\{E\left[\left(\int_s^t X_r^N(V'_N(r))\,dr\right)^2\right] : X_0^N(\mathbf{1}) \le K\right\} \le C(\varepsilon, T, K)|t-s|^{1+\theta}.$$

Hypothesis (I1) says that the mean field simplification occurs: when $V'_{N,s}$ is averaged over a macroscopic scale, the result is a constant $\gamma$ which gives the branching rate in the limiting super-Brownian motion. Hypothesis (I2) is a bound to supplement (I1), which is phrased in terms of test functions. Technically, it provides the needed control over the square function $\langle M^N(\mathbf{1})\rangle_t$. Finally, (I3) is a Kolmogorov continuity bound that will be used to obtain tightness. When the smoke clears at the end of the proof of Theorem 4.1, you will see that we could take $\theta = 1/3$.

Assuming that (H1)–(H3) and (I1)–(I3) hold, we will now use standard weak convergence arguments [see, e.g., Ethier and Kurtz (1986) or Jacod and Shiryaev (1987)] to prove weak convergence of our rescaled voter models to super-Brownian motion. Recall that $X_t(\omega) = \omega(t)$ are the coordinate variables on $\Omega_{X,D}$ and $P_N$ denotes the law of $X_t^N$. To prove tightness of the $P_N$ we use a specialized version of Jakubowski's general criterion for $D([0,\infty), E)$ when $E$ is Polish. See Theorem 3.6.4 of Dawson (1993) or Jakubowski (1986).

To state the relevant result, we begin by recalling that $\Phi \subset C_b(\mathbf{R}^d)$ is a separating class if and only if each finite measure $\mu$ on the $d$-dimensional Borel sets is uniquely determined by the values of $\mu(\phi)$ for $\phi$ in $\Phi$.

PROPOSITION 3.1. *Let $\Phi \subset C_b(\mathbf{R}^d)$ be a separating class which is closed under addition. A sequence of probabilities $P_N$ on $\Omega_{X,D}$ is tight if and only if the following conditions hold*:

(i) *For each $T, \varepsilon > 0$ there is a compact set $K_{T,\varepsilon} \subset \mathbf{R}^d$ such that*

$$\sup_N P_N\left(\sup_{t\le T} X_t(K^c_{T,\varepsilon}) > \varepsilon\right) < \varepsilon.$$

(ii) *For each $T > 0$, $\lim_{M\to\infty} \sup_N P_N\ (\sup_{t\le T}\ X_t(\mathbf{1}) > M) = 0$.*

(iii) *If $P_N^\phi(A) = P_N(X_{\bullet}(\phi) \in A)$, then for each $\phi$ in $\Phi$, $\{P_N^\phi: N \in \mathbf{N}\}$ is tight on $D = D([0,\infty), \mathbf{R})$.*

The derivation of this result from the more general results cited above is straightforward. See Theorem 3.7.1 of Dawson (1993) for the slightly simpler case where one considers measures on the one-point compactification of $\mathbf{R}^d$.

COROLLARY 3.2. *Assume $P_N$ satisfies* (i)–(iii) *of Proposition* 3.1 *with $\Phi = C_0^\infty(\mathbf{R}^d)$ and that*:

(iv) *for $\phi \in \Phi$, each limit point of $P_N^\phi$ is supported by $C([0,\infty), \mathbf{R}^d) \equiv C$.*

*Then $P_N$ is tight on $\Omega_{X,D}$ and all limit points are supported by $\Omega_{X,C} \equiv C([0,\infty), M_F(\mathbf{R}^d))$.*

PROOF. $P_N$ is tight on $\Omega_{X,D}$ by Proposition 3.1. Let $P$ be a limit point of this sequence. If $\phi \in \Phi$, then $P^\phi$ is a limit point of $P_N^\phi$ and so is supported by $C$. Let $\Phi_0$ be a countable dense subset of $\Phi$ (in the $\|\ \|_\infty$-topology). Then $P$-a.s. for all $\phi \in \Phi_0$, $X_{\bullet}(\phi)$ is continuous. As $\Phi_0$ is a separating class this shows that $X_t = X_{t-}$ for all $t \ge 0$, $P$-a.s. □

Our goal now is to verify the hypotheses (i)–(iii) of Proposition 3.1. We assume (H1)–(H3) and (I1)–(I3) throughout this section and begin with checking condition (i) in Proposition 3.1. Let $h_n: \mathbf{R}^d \to [0,1]$, be a sequence of $C^\infty$ functions such that

$$B(0,n) \subset \{x: h_n(x) = 0\} \subset \{x: h_n(x) < 1\} \subset B(0, n+1),$$

where $B(x,r)$ is the open ball in $\mathbf{R}^d$ with center $x$ and radius $r$. If we let $(h_n)_i$, $(h_n)_{ij}$ and $(h_n)_{ijk}$ denote partial derivatives of $h_n$, then

$$C_h \equiv \sup_n \sup_{1\le i,\,j,\,k\le d} \Big(\|(h_n)_i\|_\infty + \|(h_n)_{ij}\|_\infty + \|(h_n)_{ijk}\|_\infty\Big) < \infty.$$

LEMMA 3.3. For $h_n$ defined as above, $\lim_{n\to\infty} \sup_N P(\sup_{t\le T}\ X_t^N(h_n) > \varepsilon) = 0$ for any $\varepsilon, T > 0$.

PROOF. Since $h_n$ does not depend on $s$, Theorem 2.1 shows that

$$X_t^N(h_n) = X_0^N(h_n) + \int_0^t X_s^N(\mathscr{A}_N h_n)\,ds + M_t^N(h_n). \tag{3.1}$$

Clearly, we have

$$\begin{aligned}\sup_{t\le T}\left|\int_0^t X_s^N(\mathscr{A}_N h_n)\,ds\right| &\le \int_0^T X_s^N(|\mathscr{A}_N h_n - \sigma^2\Delta h_n/2|)\,ds \\ &\quad + \int_0^T X_s^N(\sigma^2|\Delta h_n|/2)\,ds.\end{aligned} \tag{3.2}$$

The first term is handled by Lemmas 2.3(a) and 2.5 which show that, uniformly in $n$,

$$\begin{aligned}&E\bigg(\int_0^T X_s^N\Big(\big|\mathscr{A}_N h_n - \sigma^2\Delta h_n/2\big|\Big)\,ds\bigg)\\ &\qquad \le \|\mathscr{A}_N h_n - \sigma^2\Delta h_n/2\|_\infty T X_0^N(\mathbf{1}) \to 0 \quad \text{as } N\to\infty,\end{aligned}$$

As the left side trivially approaches 0 as $n\to\infty$ for each fixed value of $N$ by dominated convergence and Lemma 2.3(a), it follows immediately that

$$\lim_{n\to\infty}\sup_N E\bigg(\int_0^T X_s^N(|\mathscr{A}_N h_n - \sigma^2\Delta h_n/2|)\,ds\bigg) = 0. \tag{3.3}$$

By Lemma 2.3(a), using Chebyshev's inequality for the last inequality, the mean of the second term on the right side of (3.2) is at most

$$\begin{aligned}CE\bigg(\int_0^T X_s^N(B(0,n)^c)\,ds\bigg) &= CE\bigg(\int_0^T\int P(|B_s^{x,N}|\ge n)\,X_0^N(dx)\,ds\bigg)\\ &\le CT\Big[X_0^N(B(0,n/2)^c)\\ &\qquad + X_0^N(\mathbf{1})\sup\nolimits_{s\le T} P(|B_s^{0,N}|\ge n/2)\Big]\\ &\le CT\Big[X_0^N(B(0,n/2)^c) + X_0^N(\mathbf{1})CTn^{-2}\Big],\end{aligned} \tag{3.4}$$

which approaches 0 uniformly in $N$ as $n\to\infty$ by (H2). Combine this with (3.2) and (3.3) to see that

$$\lim_{n\to\infty}\sup_N E\bigg(\sup_{t\le T}\left|\int_0^t X_s^N(\mathscr{A}_N h_n)\,ds\right|\bigg) = 0 \quad \text{for all } T>0. \tag{3.5}$$

Apply (3.1) and (3.5) to see that for some $\eta_n\to 0$,

$$E\Big(\big|M_T^N(h_n)\big|\Big) \le E(X_T^N(h_n)) + X_0^N(h_n) + \eta_n = X_0^N(P_T^N h_n + h_n) + \eta_n,$$

the last by Lemma 2.3(a). To bound the right-hand side, we note that $B(0,n)\subset \{x: h_n(x)=0\}$ and $h_n\le 1$, and thus

$$X_0^N(P_T^N h_n) \le X_0^N(B(0,n/2)^c) + X_0^N(\mathbf{1})P\Big(\big|B_s^{0,N}\big| > n/2\Big) \to 0 \quad \text{as } n\to 0$$

uniformly in $N$. Since $X_0^N(h_n) \le X_0^N(B(0,n)^c)$, we can conclude from (H2) that

$$\lim_{n\to\infty} \sup_N X_0^N(P_T^N h_n + h_n) = 0.$$

The weak $L^1$ inequality for martingales now gives

$$\lim_{n\to\infty} \sup_N P\left( \sup_{t\le T} \left|M_t^N(h_n)\right| > \varepsilon \right) = 0 \quad \text{for all } \varepsilon > 0. \tag{3.6}$$

Since (3.6), (3.5) and (H2) control the three terms in (3.1), the proof of Lemma 3.3 [and condition (i) of Proposition 3.1] is now complete. □

To prove (ii) in Proposition 3.1, we begin by observing that (2.5) gives $X_t^N(\mathbf{1}) = X_0^N(\mathbf{1}) + M_t^N(\mathbf{1})$ while (2.7) tells us that

$$\langle X^N(\mathbf{1})\rangle_t = \langle M^N(\mathbf{1})\rangle_t = 2\int_0^t X_s^N(V'_{N,s})\,ds.$$

If we use (I2) and apply Doob's $L^2$-inequality to the martingale $M_t^N(\mathbf{1})$ we get

$$E\left( \sup_{s\le T} M_s^N(\mathbf{1})^2 \right) \le C_T X_0^N(\mathbf{1}).$$

From this it follows easily that, for a finite constant $C_T$,

$$E\left( \sup_{s\le T} X_s^N(\mathbf{1})^2 \right) \le C_T(X_0^N(\mathbf{1}) + X_0^N(\mathbf{1})^2), \tag{3.7}$$

and using (H2) we have the desired result by Chebyshev's inequality.

Next we verify (iii) in Proposition 3.1 and (iv) in Corollary 3.2.

LEMMA 3.4. *If $\phi \in C_0^\infty(\mathbf{R}^d)$, then $\{P_N^\phi : N \in \mathbf{N}\}$ is tight in $D$ and all limit points are in $C([0,\infty),\mathbf{R})$.*

PROOF. To do this, we begin by using (2.5) to write

$$X_t^N(\phi) = X_0^N(\phi) + \int_0^t X_s^N(\mathscr{A}_N \phi_s)\,ds + M_t^N(\phi). \tag{3.8}$$

The first term on the right is constant in $t$. Our approach will be to show that the other two terms are also tight. If $t < u \le T$, then the uniform convergence of $\mathscr{A}_N\phi$ to $\sigma^2\Delta\phi/2$ proved in Lemma 2.5 implies there is a finite constant $C_\phi$ such that

$$E\left[ \left( \int_t^u |X_s^N(\mathscr{A}_N\phi)|\,ds \right)^2 \right] \le C_\phi E\left[ \left( \int_t^u X_s^N(\mathbf{1})\,ds \right)^2 \right].$$

Inequality (3.7) shows that if $X_0^N(\mathbf{1}) \le K$, then the right side above is not larger than

$$C_\phi E\left[ \sup_{s\le T} X_s^N(\mathbf{1})^2 \right](u-t)^2 \le C_{\phi,T}(K + K^2)(u-t)^2.$$

Let $A_N(t) = \int_0^t X_s^N(\mathscr{A}_N\phi)\,ds$. The above bound and Proposition VI.3.26 in Jacod and Shiryaev (1987) imply that $A_N$ defines a tight sequence of probabilities on $C([0,\infty),\mathbf{R})$.

Turning now to $M_t^N(\phi)$, Theorem 2.1 implies

$$\langle M^N(\phi)\rangle_t = \int_0^t [2X_s^N(\phi_s^2 V'_{N,s}) + \varepsilon_s^N(\phi)]\,ds.$$

From this we see that for $0 < \varepsilon \le s < t \le T$, $t-s \le 1$,

$$\begin{aligned} &E\Big(\{\langle M^N(\phi)\rangle_t - \langle M^N(\phi)\rangle_s\}^2\Big) \\ &\quad \le C\left\{\|\phi\|_\infty^4 E\left[\left(\int_s^t X_r^N(V'_{N,r})\,dr\right)^2\right] + E\Big(\sup_{r\le T}\varepsilon_r^N(\phi)^2\Big)(t-s)^2\right\}, \end{aligned} \tag{3.9}$$

where $C$ is a finite constant. Using (I3) and Theorem 2.1, we conclude that if $N$ is large and $X_0^N(\mathbf{1}) \le K$, the right side of (3.9) is at most

$$C\Big\{\|\phi\|_\infty^4 C(\varepsilon,T,K)(t-s)^{1+\theta} + (t-s)^2\Big\} \le C(\varepsilon,T,K,\phi)(t-s)^{1+\theta}$$

for $\varepsilon \le s < t \le T$, since we have assumed in (I3) that $0 < \theta \le 1$.

Now Theorem 2.1 and (I2) show that

$$\begin{aligned} E(\langle M^N(\phi)\rangle_T) &\le 2\|\phi\|_\infty^2 E\left(\int_0^T X_s^N(V'_{N,s})\,ds\right) + TE\Big(\sup_{s\le T}\varepsilon_s^N(\phi)\Big) \\ &\le 2\|\phi\|_\infty^2 C_T X_0^N(\mathbf{1}) + TE\Big(\sup_{s\le T}\varepsilon_s^N(\phi)\Big). \end{aligned}$$

The right side is bounded uniformly in $N$ by (H2) and Theorem 2.1 and approaches 0 as $T \downarrow 0$ uniformly in $N$ by Theorem 2.1 and (I2).

It is now an elementary exercise to use the above two inequalities to show that $\langle M^N(\phi)\rangle_\cdot$ is tight in $C([0,\infty),\mathbf{R})$, for example, by using Proposition VI.3.26 in Jacod and Shiryaev (1987) again. Recall also that $\sup_{t\le T} |\Delta M^N(\phi)(t)| \le \|\phi\|_\infty (N')^{-1}$. Theorem VI.4.13 and Proposition VI.3.26 of Jacod and Shiryaev (1987) show that $\{M^N(\phi): N \in \mathbf{N}\}$ is tight in $D$ and all limit points are supported by $C([0,\infty),\mathbf{R})$. Using (3.8) now with the above results, we obtain the desired conclusion. □

At this point we have proved that all the hypotheses of Corollary 3.2 are satisfied so we are ready to prove our main result. Recall from Section 1 that $P_{X_0}^{b,\sigma^2}$ denotes the law of super-Brownian motion on $\Omega_{X,D}$ or $\Omega_{X,C}$, starting at $X_0$, and with branching rate $b$ and diffusion coefficient $\sigma^2$.

THEOREM 3.5. *Assume* (H1)–(H3) *and* (I1)–(I3). *Then* $P_N \Rightarrow P_{X_0}^{2\gamma,\sigma^2}$.

PROOF. Corollary 3.2 implies that $\{P_N\}$ is tight and all limit points are supported by $\Omega_{X,C}$. Let $P$ be any such limit point and let $\phi \in C_0^\infty(\mathbf{R}^d)$. By Skorohod's theorem [Theorem 3.1.8 in Ethier and Kurtz (1986)] we may assume there is an $X$ with law $P$ and a subsequence $\{N_k\}$ such that

$$X^{N_k} \to X \quad \text{a.s. in } \Omega_{X,D}. \tag{3.10}$$

Lemma 2.5 and part (iii) of Theorem 2.1 show that for all $T > 0$,

$$\lim_{k\to\infty} \sup_{t\le T} \left| \int_0^t X_s^{N_k}(\mathscr{A}_{N_k}\phi)\,ds - \int_0^t X_s(\sigma^2\Delta\phi/2)\,ds \right| = 0 \quad \text{a.s.} \tag{3.11}$$

Let $M_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s(\sigma^2\Delta\phi/2)\,ds$. Then (3.10) and (3.11) together with the identity in (i) of Theorem 2.1 imply that for all $T > 0$,

$$\lim_{k\to\infty} \sup_{t\le T} \left| M_t^{N_k}(\phi) - M_t(\phi) \right| = 0 \quad \text{a.s.} \tag{3.12}$$

Here we use the fact that $M_\cdot(\phi)$ is continuous to derive uniform convergence on compacts from convergence in $D$.

Lemma 2.4 implies that

$$E\langle M^N(\phi)\rangle_t^2 = E\left(\int_0^t \left[2X_s^N(\phi^2 V'_{N,s}) + \varepsilon_s^N(\phi)\right] ds\right)^2,$$

where $|\varepsilon_s^N(\phi)| \le C'_\phi X_s^N(\mathbf{1})\sqrt{N}/N'$. It follows from (I1), Theorem 2.1(ii), and (3.7), that for all $T > 0$, $\sup_N E\left(\langle M^N(\phi)\rangle_T^2\right) < \infty$. Using Burkholder's inequality in (2.9) and $|\Delta M^N(\phi)(t)| \le \|\phi\|_\infty (N')^{-1}$, we obtain

$$\sup_N E\left(\sup_{t\le T} |M_t^N(\phi)|^4\right) < \infty \quad \text{for all } T > 0. \tag{3.13}$$

Fix $0 \le t_1 < t_2 < \cdots < t_n \le s < t$ and test functions $h_i\colon \mathscr{M}_F(\mathbf{R}^d) \to \mathbf{R}$ that are bounded and continuous for $1 \le i \le n$. Now (3.10), (3.12), (3.13) and dominated convergence imply that

$$E\left((M_t(\phi) - M_s(\phi))\prod_1^n h_i(X_{t_i})\right)$$
$$= \lim_{k\to\infty} E\left((M_t^{N_k}(\phi) - M_s^{N_k}(\phi))\prod_1^n h_i(X_{t_i}^{N_k})\right) = 0,$$

the last from Theorem 2.1. Therefore under $P$, $M_t(\phi)$ is a continuous $\mathscr{F}_t^X$-martingale where $\mathscr{F}_t^X$ is the canonical right-continuous filtration generated by $X$. Also, (3.10), (3.12), (3.13) and (3.7) imply

$$E\left(\left(M_t(\phi)^2 - M_s(\phi)^2 - \int_s^t X_r(2\gamma\phi^2)\,dr\right)\prod_1^n h_i(X_{t_i})\right)$$
$$= \lim_{k\to\infty} E\left(\left(M_t^{N_k}(\phi)^2 - M_s^{N_k}(\phi)^2 - \int_s^t X_r^{N_k}(2\gamma\phi^2)\,dr\right)\prod_1^n h_i(X_{t_i}^{N_k})\right).$$

Lemma 2.4 and (I1) show that the above equals

$$\lim_{k\to\infty} E\left(\left(M_t^{N_k}(\phi)^2 - M_s^{N_k}(\phi)^2 - (\langle M^{N_k}(\phi)\rangle_t - \langle M^{N_k}(\phi)\rangle_s)\right)\prod_1^n h_i(X_{t_i}^{N_k})\right),$$

which is 0 by Theorem 2.1. This shows that $\langle M(\phi)\rangle_t = \int_0^t X_s(2\gamma\phi^2)\,ds$ for all $t \geq 0$, $P$-a.s. We have proved that $X$ satisfies the martingale problem $(MP)_{X_0}^{2\gamma,\sigma^2}$ (see Section 1) which characterizes the law of super-Brownian motion (see the Appendix). Therefore the law of $X$, $P$, equals $P_{X_0}^{2\gamma,\sigma^2}$. As $P$ is an arbitrary limit point, the result follows. □

**4. Analysis of the square function.** In this section we will show that in the presence of (H1)–(H3), (K1)–(K4) imply the intermediate conditions (I1)–(I3). To do this, we will introduce a second set of intermediate conditions (J1)–(J2), but for that we need some notation. For $\phi$ in $C_0^\infty(\mathbf{R}^d)$ and $\gamma > 0$, define

$$\varepsilon_{K,\phi}^{N,\gamma}(t) = \sup\left\{\left|E^{\xi_0^N}[X_t^N(\{V'_{N,t}-\gamma\}\phi^2)]\right| : X_0^N(\mathbf{1}) \leq K\right\}.$$

(J1) There is a $\gamma > 0$ such that for all $\phi \in C_0^\infty(\mathbf{R}^d)$, $t > 0$ and $K < \infty$,

$$\lim_{N\to\infty} \varepsilon_{K,\phi}^{N,\gamma}(t) = 0.$$

(J2) There is a $p > 1$ and function $g: [0,\infty) \mapsto [1,\infty]$ which is bounded on compact subintervals of $(0,\infty)$, such that for all $T > 0$, $\int_0^T g(s)^p\,ds < \infty$, and for each $T > 0$ there exists a finite $C = C_T$ such that for all $s \in [0,T]$ and $N \in \mathbf{N}$,

$$\text{(a)}\qquad EX_s^N(V'_{N,s}) \leq g(s)X_0^N(\mathbf{1}),$$

$$\text{(b)}\qquad EX_s^N(\mathbf{1})^3 \leq C_T(X_0^N(\mathbf{1})^3 + 1),$$

$$\text{(c)}\qquad E\left[X_s^N(V'_{N,s})X_s^N(\mathbf{1})^2\right] \leq g(s)(X_0^N(\mathbf{1}) + 1).$$

Obviously (J1) will be the key to the proof of (I1), but note that (J1) only says that the mean of $V'_{N,s}$ is close to $\gamma$ while (I1) asserts convergence of $V'_{N,s}$ to $\gamma$ in an $L^2$ sense. The conditions in (J2) are boundedness conditions that will be needed to upgrade (J1) to (I1) and establish the earlier regularity conditions (I2) and (I3). It will turn out (see the proof of Lemma 4.5 below) that we can take $g(s) = C_s(1+s^{-1/2})$ where $s \to C_s$ is increasing, and hence we can choose $p = 3/2$. Note that we have assumed $g(s) \geq 1$, so $g(s)$ can be used to absorb constants.

THEOREM 4.1. *Assume* (H1)–(H3). *If* (J1)–(J2) *hold, then* (I1)–(I3) *hold.*

PROOF OF (I1). Let $\phi \in C_0^\infty(\mathbf{R}^d)$ and set

$$\phi_{N,s}(x) = \{V'_{N,s}(x) - \gamma\}\phi^2(x). \tag{4.1}$$

By expanding the expectation in (I1), we obtain

$$E\left[\left(\int_0^T X_s^N(\phi_{N,s})\,ds\right)^2\right] = 2\int_0^T \int_s^T E\Big(X_s^N(\phi_{N,s})X_t^N(\phi_{N,t})\Big)\,dt\,ds.$$

By the Markov property, for $s < t$,

$$E\Big(X_s^N(\phi_{N,s})X_t^N(\phi_{N,t})\Big) = E\Big[X_s^N(\phi_{N,s})\,E^{\xi_s^N}\Big(X_{t-s}^N(\phi_{N,t-s})\Big)\Big].$$

Thus,

$$\begin{aligned} E&\left[\left(\int_0^T X_s^N(\phi_{N,s})\,ds\right)^2\right] \\ &= 2E\left[\int_0^T X_s^N(\phi_{N,s})\int_s^T E^{\xi_s^N}\Big(X_{t-s}^N(\phi_{N,t-s})\Big)\,dt\,ds\right]. \end{aligned} \tag{4.2}$$

Using the definition of $\phi_{N,s}$ and (J2.a), we have for any $s > 0$,

$$\begin{aligned} E^{\xi_0^N}\Big(\Big|X_s^N(\phi_{N,s})\Big|\Big) &\le E^{\xi_0^N}\Big(X_s^N(V'_{N,s}\phi^2) + \gamma\|\phi\|_\infty^2 X_s^N(\mathbf{1})\Big) \\ &\le \|\phi\|_\infty^2 g(s)X_0^N(\mathbf{1}) + \gamma\|\phi\|_\infty^2 X_0^N(\mathbf{1}) \\ &\le (1+\gamma)g(s)\|\phi\|_\infty^2 X_0^N(\mathbf{1}), \end{aligned}$$

where in the last step we have used $g(s) \ge 1$. Using the definition of $\varepsilon_{K,\phi}^{N,\gamma}(s)$ now it follows that

$$\varepsilon_{K,\phi}^{N,\gamma}(s) \le (1+\gamma)g(s)\|\phi\|_\infty^2 K. \tag{4.3}$$

Let $\Omega_K(s) = \{X_s^N(\mathbf{1}) \le K\}$. On $\Omega_K(s)$ we have

$$|E^{\xi_s^N}(X_{t-s}^N(\phi_{N,t-s}))| \le \varepsilon_{K,\phi}^{N,\gamma}(t-s) \le C_{\phi,K}\,g(t-s), \tag{4.4}$$

where $C_{\phi,K} = (1+\gamma)\|\phi\|_\infty^2 K$. Thus, by (H2), (J1), the dominated convergence theorem and the fact that $\int_0^T g(s)\,ds < \infty$,

$$\begin{aligned} 2E&\Big(\int_0^T \big|X_s^N(\phi_{N,s})\big|1_{\Omega_K(s)}\int_s^T \big|E^{\xi_s^N}(X_{t-s}^N(\phi_{N,t-s}))\big|\,dt\,ds\Big) \\ &\le 2C_{\phi,K}X_0^N(\mathbf{1})\int_0^T g(s)\,ds\int_0^T \varepsilon_{K,\phi}^{N,\gamma}(t)\,dt \to 0 \end{aligned} \tag{4.5}$$

as $N \to \infty$.

To handle the contribution from $\Omega_K^c(s)$, we note that using the definition of $\phi_{N,s}$ then (J2.a) and part (a) of Lemma 2.3, we have that for all $s \le T$,

$$\begin{aligned}\int_s^T \big|E^{\xi_s^N} X_{t-s}^N(\phi_{N,t-s})\big|\,dt &\le (1+\gamma)\|\phi\|_\infty^2 \int_s^T E^{\xi_s^N}\big(X_{t-s}^N(V'_{N,t-s}) + X_{t-s}^N(\mathbf{1})\big)\,dt \\ &\le (1+\gamma)\|\phi^2\|_\infty X_s^N(\mathbf{1}) \int_s^T (g(t-s)+1)\,dt \\ &\le C_{T,\phi} X_s^N(\mathbf{1}),\end{aligned}$$

where $C_{T,\phi} = (1+\gamma)\|\phi\|_\infty^2 \int_0^T g(s)\,ds$. On account of this, and the definition of $\phi_{N,s}$,

$$\begin{aligned}&2E\int_0^T |X_s^N(\phi_{N,s})| 1_{\Omega_K^c(s)} \int_s^T \big|E^{\xi_s^N}[X_{t-s}^N(\phi_{N,t-s})]\big|\,dt \\ &\le 2C_{T,\phi}\int_0^T E\Big[\big|X_s^N(\phi_{N,s})\big| X_s^N(\mathbf{1}) 1_{\Omega_K^c(s)}\Big] \qquad (4.6)\\ &\le 2C_{T,\phi}(1+\gamma)\|\phi^2\|_\infty \int_0^T E\big[\{X_s^N(V'_{N,s}) + X_s^N(\mathbf{1})\} X_s^N(\mathbf{1}) 1_{\Omega_K^c(s)}\big]\,ds.\end{aligned}$$

Using the inequality $E|UV|1_{\{|V|>K\}} \le E(|U|V^2)/K$, and then (J2.b) and (J2.c) it follows that for $s \in [0,T]$,

$$E\big[\{X_s^N(V'_{N,s}) + X_s^N(\mathbf{1})\} X_s^N(\mathbf{1}) 1_{\Omega_K^c(s)}\big] \le \frac{C}{K}(X_0^N(\mathbf{1})^3 + 1) g(s),$$

for a finite constant $C$. Using this inequality in (4.6) we obtain

$$\begin{aligned}&2E\Big(\int_0^T \big|X_s^N(\phi_{N,s})\big| 1_{\Omega_K^c(s)} \int_s^T \big|E^{\xi_s^N}[X_{t-s}^N(\phi_{N,t-s})]\big|\,dt\,ds\Big) \\ &\qquad \le \frac{2C_{T,\phi}^2 C}{K}(X_0^N(\mathbf{1})^3 + 1).\end{aligned} \qquad (4.7)$$

By combining (4.5) and (4.7), we obtain (I1). □

PROOF OF (I2). This follows immediately from (J2.a), with $C_T = \int_0^T g(s)\,ds$. □

PROOF OF (I3). We fix $0 < \varepsilon < T < \infty$, and consider $s, t$ such that $\varepsilon \le s < t \le T$. The first step is to note that

$$E\Big(\int_s^t X_r^N(V'_{N,r})\,dr\Big)^2 = 2\int_s^t \int_r^t E[X_r^N(V'_{N,r}) X_{r'}^N(V'_{N,r'})]\,dr'\,dr.$$

Let $\varepsilon \le r \le r' \le T$ and note $X_r^N(\mathbf{1}) \le 1 + X_r^N(\mathbf{1})^2$. By the Markov property, and all three bounds in (J2) we get

$$E\Big[X_r^N(V'_{N,r}) X_{r'}^N(V'_{N,r'})\Big] = E\Big[X_r^N(V'_{N,r}) E^{\xi_r^N} X_{r'-r}^N(V'_{N,r'-r})\Big]$$

$$\le E\Big[X_r^N(V'_{N,r})X_r^N(\mathbf{1})g(r'-r)\Big]$$
$$\le (X_0^N(\mathbf{1})^3+1)g(r)g(r'-r)$$
$$\le C_{\varepsilon,T}(X_0^N(\mathbf{1})^3+1)g(r'-r)$$

for a finite constant $C_{\varepsilon,T}$, where we have used the boundedness of $g$ on $[\varepsilon, T]$. A little calculus and Hölder's inequality show that if $p$ is the power in (J2) and $1/q = 1-1/p$, then

$$\int_s^t \int_r^t g(r'-r)\,dr'dr = \int_0^{t-s} g(u)(t-s-u)\,du$$
$$\le \left(\int_0^{t-s} g(u)^p\,du\right)^{1/p}\left(\frac{|t-s|^{q+1}}{q+1}\right)^{1/q}.$$

Consequently, for $\theta = 1/q$, we can choose a $C(\varepsilon, T, K)$ so that (I3) holds. This completes the proof of Theorem 4.1. In closing we would like to note that if $p = 3/2$ as advertised in the discussion of (J2) above, then $q = 3$ and $\theta = 1/3$ as promised in the introduction. □

The second half of closing the gap between the $K$'s and the $I$'s is shown in the following.

THEOREM 4.2. *Assume* (H1)–(H3). *If* (K1)–(K4) *hold*, *then* (J1)–(J2) *hold*.

The first half of that is to show the following lemma.

LEMMA 4.3. *Assume* (H1)–(H3). *If* (K1)–(K3) *hold then* (J1) *holds*.

Note that the technical condition (K4) is not needed for the proof of (J1). It will come up several times in the proof of (J2).

PROOF. Fix $t > 0$ and $\phi \in C_0^\infty(\mathbf{R}^d)$. Recalling the definition of $V'_N$,

$$E\big(X_t^N(V'_{N,t}\phi^2)\big) = \frac{1}{N'}\sum_x \phi(x)^2 \sum_y p_N(y-x)E\big[\xi_t^N(x)(1-\xi_t^N(y))\big]N/N'.$$

Here for the first time we will use the usual particle system duality between the voter model and coalescing random walks. The reader can find this described in Griffeath (1978) or in Section 3 of Durrett (1995). Intuitively, if we follow the sources of the opinions at $x$ and $y$ at time $t$ backwards in time, then the result is a pair of dual random walks $\hat{B}_t^{x,N}$ and $\hat{B}_t^{y,N}$ which: (i) individually behave like $B_t^{x,N}$ and $B_t^{y,N}$, (ii) are independent until they collide, and (iii) stay together after they collide.

To say this more formally, let $\tau^N(x,y) = \inf\{s\colon B_s^{x,N} = B_s^{y,N}\}$, and let $\hat{B}_t^{x,N} = B_t^{x,N}$ for all $t$ and

$$\hat{B}_t^{y,N} = \begin{cases} B_t^{y,N}, & t \le \tau_{x,y}^N,\\ B_t^{x,N}, & t \ge \tau_{x,y}^N.\end{cases}$$

Duality implies that $P(\xi_t^N(x)=1,\ \xi_t^N(y)=0) = P(\xi_0(\hat{B}_t^{x,N})=1,\ \xi_0(\hat{B}_t^{y,N})=0)$, so

$$E\big(X_t^N(V'_{N,t}\phi^2)\big) = \frac{1}{N'}\sum_x \phi(x)^2 \sum_y p_N(y-x) \times P\big[\xi_0^N(B_t^{x,N})=1,\ \xi_0^N(B_t^{y,N})=0,\ \tau^N(x,y)>t\big]\frac{N}{N'}.$$

To estimate the last sum we will write it as $\Sigma_1^N - \Sigma_2^N$, where

$$\Sigma_1^N = \frac{1}{N'}\sum_x \phi(x)^2 \sum_y p_N(y-x) P\Big[\xi_0^N(B_t^{x,N})=1,\ \tau^N(x,\ y)>t\Big]N/N',$$

$$\Sigma_2^N = \frac{1}{N'}\sum_x \phi(x)^2 \sum_y p_N(y-x) \times P\Big[\xi_0^N(B_t^{x,N})=1,\ \xi_0^N(B_t^{y,N})=1,\ \tau^N(x,\ y)>t\Big]N/N'.$$

We first show that $\Sigma_2^N \to 0$ uniformly in $X_0^N(\mathbf{1}) \le K$ as $N\to\infty$. Summing over the possible values of $w = B_t^x$ and $z = B_t^y$ and setting $y - x = e$, we get

$$\Sigma_2^N = \frac{1}{(N')^2}\sum_w\sum_z \xi_0^N(w)\xi_0^N(z)\sum_e p_N(e)\sum_x \phi(x)^2 \times P\Big[B_t^{x,N}=w,\ B_t^{x+e,N}=z,\ \tau^N(x,x+e)>t\Big]N. \tag{4.8}$$

To explain our arithmetic below, we first note that $(1/(N')^2)\sum_w\sum_z \xi_0^N(w)\xi_0^N(z) = X_0^N(\mathbf{1})^2$. Next, it is easy to see that

$$\begin{aligned}
&\sum_e p_N(e)\sum_x \phi(x)^2\, P\Big[B_t^{x,N}=w,\ B_t^{x+e,N}=z,\ \tau^N(x,\ x+e)>t\Big]N\\
&\quad\le \sum_e p_N(e)\|\phi\|_\infty^2 \sum_x P\Big[B_t^{0,N}=w-x,\ B_t^{e,N}=z-x,\ \tau^N(0,e)>t\Big]N\\
&\quad= \sum_e p_N(e)\|\phi\|_\infty^2 P\Big[B_t^{0,N}-B_t^{e,N}=w-z,\ \tau^N(0,\ e)>t\Big]N.
\end{aligned}$$

By the Markov property, we have

$$\begin{aligned}
&\sum_e p_N(e)P\Big[B_t^{0,N}-B_t^{e,N}=w-z,\ \tau^N(0,e)>t\Big]N\\
&\quad\le \sum_e p_N(e)E\Big(1(\tau^N(0,\ e)>t/2)P\Big[B_t^{0,N}-B_t^{e,N}=w-z\,|\,B_{t/2}^{0,N},\,B_{t/2}^{e,N}\Big]\Big)N\\
&\quad\le \frac{N}{N'}P(\tau^{*,N}>t/2)N'\sup_x P(B_t^{0,N}=x) \equiv \eta_N(t).
\end{aligned}$$

Combining these estimates we obtain $\Sigma_2^N \le X_0^N(\mathbf{1})^2\|\phi\|_\infty^2\eta_N(t)$. Since (K1) and (K3) imply that $\eta_N(t)\to 0$, we conclude that

(4.9) $$\sup_{X_0^N(\mathbf{1})\le K} \Sigma_2^N \le K^2\|\phi\|_\infty^2\eta_N(t)\to 0 \qquad \text{as } N\to\infty.$$

Turning now to $\Sigma_1^N$, we can again sum over all the possible values of $w = B_t^{x,N}$ to write

$$\Sigma_1^N = \frac{1}{N'}\sum_w \xi_0^N(w)\sum_e p_N(e)\sum_x \phi(x)^2 P\Big(B_t^{x,N}=w,\ \tau^N(x,x+e)>t\Big)N/N'.$$

Using translation, reflection symmetry, and then translation again, we have

$$\begin{aligned} P\Big(B_t^{x,N}=w,\ \tau^N(x,x+e)>t\Big) &= P\Big(B_t^{0,N}=w-x,\ \tau^N(0,e)>t\Big)\\ &= P\Big(B_t^{0,N}=x-w,\ \tau^N(0,-e)>t\Big)\\ &= P\Big(B_t^{w,N}=x,\ \tau^N(w,w-e)>t\Big). \end{aligned}$$

Using symmetry again we get

$$\Sigma_1^N = \frac{1}{N'}\sum_w \xi_0^N(w)\sum_e p_N(e)\sum_x \phi(x)^2 P\Big(B_t^{w,N}=x,\ \tau^N(w,w+e)>t\Big)N/N'.$$

We will write this as $\Sigma_3^N-\Sigma_4^N$, where [recall $\varepsilon_N$ in (K2)]

$$\Sigma_3^N = \frac{1}{N'}\sum_w \xi_0^N(w)\sum_e p_N(e)E\Big[\phi(B_t^{w,N})^2 1\{\tau^N(w,w+e)>\varepsilon_N t\}\Big]N/N',$$

$$\Sigma_4^N = \frac{1}{N'}\sum_w \xi_0^N(w)\sum_e p_N(e)E\Big[\phi(B_t^{w,N})^2 1\{\tau^N(w,w+e)\in(\varepsilon_N t, t]\}\Big]N/N'.$$

To bound $\Sigma_4^N$, we note that

$$\begin{aligned} \Sigma_4^N &\le \frac{\|\phi\|_\infty^2}{N'}\sum_w \xi_0^N(w)\sum_e p_N(e)P\Big(\tau^N(0,e)\in(\varepsilon_N t,t]\Big)N/N'\\ &= \|\phi\|_\infty^2 X_0^N(\mathbf{1})\frac{N}{N'}P(\tau^{*,N}/2\in(\varepsilon_N t,t]), \end{aligned}$$

where the factor of 2 arises because $B_t^{0,N}-B_t^{e,N}$ jumps at rate $2N$. Condition (K2.b) now gives

(4.10) $$\sup_{X_0^N(\mathbf{1})\le K} \Sigma_4^N\to 0 \qquad \text{as } N\to\infty.$$

Combining (4.9), (4.10) and Lemma 2.3(a) will show that the proof of Lemma 4.3 will be complete once we establish

(4.11) $$\sup_{X_0^N(\mathbf{1})\le K} \big|\Sigma_3^N-\gamma X_0^N(P_t^N(\phi^2))\big|\to 0 \quad \text{as } N\to\infty.$$

The idea behind (4.11) is that $\phi^2(B_t^{w,N})$ and $1\{\tau^N(w, w+e) > \varepsilon_N t\}$ are almost independent, and thus

$$\begin{aligned}\Sigma_3^N &= \frac{1}{N'}\sum_w \xi_0^N(w)\sum_e p_N(e)E\Big[\phi^2(B_t^{w,N})1\{\tau^N(w, w+e) > \varepsilon_N t\}\Big]\frac{N}{N'}\\ &\approx \frac{1}{N'}\sum_w \xi_0^N(w)E[\phi^2(B_t^{w,N})]\sum_e p_N(e)P(\tau^N(0,e) > \varepsilon_N t)\frac{N}{N'}\\ &= X_0^N(P_t^N(\phi^2))\frac{N}{N'}P(\tau^{*,N} > 2\varepsilon_N t) \approx \gamma X_0^N(P_t^N(\phi^2)),\end{aligned}$$

by (K1) and (K2)(b).

PROOF OF (4.11). Recall that $P_s^N f(x) = Ef(B_s^{x,N})$. The Markov property implies that

$$\begin{aligned}&E\Big[\phi^2(B_t^{w,N})1\{\tau^N(w, w+e) > \varepsilon_N t\}\Big]\\ &\quad= E\Big[1\{\tau^N(w, w+e) > \varepsilon_N t\}P^N_{(1-\varepsilon_N)t}\phi^2(B^{w,N}_{\varepsilon_N t})\Big].\end{aligned}$$

Our goal is to show that a small error results if we replace $P^N_{(1-\varepsilon_N)t}\phi^2(B^{w,N}_{\varepsilon_N t})$ by $P_t^N\phi^2(w)$. By (K2)(a) we may choose $\delta_N \downarrow 0$ such that

$$\lim_{N\to\infty}\frac{N}{N'}P\Big(|B^{0,N}_{t\varepsilon_N}| > \delta_N\Big) = 0. \tag{4.12}$$

Let $B_t$ be a $d$-dimensional Brownian motion starting at 0, with covariance matrix $\sigma^2 I$, and let $P_t$ denote its semigroup. As $\phi^2$ is Lipschitz we may choose $C_\phi$ so that $|\phi^2(x)-\phi^2(y)| \le C_\phi|x-y|$. If $t_N \to t$ and $z_N \to z \in \mathbf{R}^d$ as $N \to \infty$, then

$$\begin{aligned}&|P^N_{t_N}(\phi^2)(z_N) - P_t(\phi^2)(z)|\\ &\quad\le \big|E(\phi^2(z_N + B^{0,N}_{t_N}) - \phi^2(z + B^{0,N}_{t_N}))\big|\\ &\qquad+ \big|E(\phi^2(z + B^{0,N}_{t_N})) - E(\phi^2(z + B_t))\big|\\ &\quad\le C_\phi|z - z_N| + \big|E(\phi^2(z + B^{0,N}_{t_N})) - E(\phi^2(z + B_t))\big| \to 0.\end{aligned} \tag{4.13}$$

For the convergence in the last line, see Theorem VIII.3.33 of Jacod and Shiryaev (1987) and use (H1) to verify the hypotheses of that result.

If $t_N \to t$, $|z_N| \to \infty$, and the support of $\phi$ is contained in $B(0, R)$, then

$$\big|P^N_{t_N}\phi^2(z_N)\big| \le \|\phi\|_\infty^2 P\big(|z_N + B^{0,N}_{t_N}| \le R\big) \le \|\phi\|_\infty^2 P(|B^{0,N}_{t_N}| > |z_N| - R).$$

Using (H1)(a), (H1)(c) and Chebyshev's inequality, it follows that

$$\big|P^N_{t_N}\phi^2(z_N)\big| \le \|\phi\|_\infty^2\frac{Ct_N}{(|z_N| - R)^2} \to 0 \tag{4.14}$$

as $N \to \infty$. By (4.13) and (4.14),

$$\kappa_N(t) \equiv \sup\left\{\left|P_s^N(\phi^2)(w) - P_t^N(\phi^2)(z)\right| : |w-z| \le \delta_N,\ |s-t| \le \varepsilon_N t\right\} \to 0 \tag{4.15}$$

as $N \to 0$.

With (4.15) in hand the rest is straightforward. To begin, we use the Markov property to see that

$$\begin{aligned} \Sigma_3^N - \gamma X_0^N(P_t^N \phi^2) &= \frac{1}{N'} \sum_w \xi_0^N(w) \sum_e p_N(e) \\ &\quad \times E\left[1(\tau^N(w, w+e) > \varepsilon_N t) \frac{N}{N'} (P_{(1-\varepsilon_N)t}^N \phi^2)(B_{\varepsilon_N t}^{w,N})\right] \\ &\quad - \frac{\gamma}{N'} \sum_w \xi_0^N(w)(P_t^N \phi^2)(w). \end{aligned} \tag{4.16}$$

Next, we rewrite the second term on the right side above as

$$\begin{aligned} &- \frac{1}{N'} \sum_w \xi_0^N(w) \sum_e p_N(e) E[1\{\tau^N(w, w+e) > \varepsilon_N t\}] \frac{N}{N'} P_t^N \phi^2(w) \\ &+ \frac{1}{N'} \sum_w \xi_0^N(w) \left[\sum_e p_N(e) \left\{P(\tau^N(w, w+e) > \varepsilon_N t) \frac{N}{N'} - \gamma\right\}\right] P_t^N \phi^2(w). \end{aligned}$$

Substituting this into (4.16) we obtain

$$\begin{aligned} &\left| \Sigma_3^N - \gamma X_0^N(P_t^N \phi^2)\right| \\ &\quad \le \frac{1}{N'} \sum_w \xi_0^N(w) P(|B_{\varepsilon_N t}^{w,N} - w| > \delta_N) \frac{N}{N'} 2\|\phi\|_\infty^2 \\ &\quad + \frac{1}{N'} \sum_w \xi_0^N(w) \sum_e p_N(e) E\Big(1\{\tau^N(w, w+e) > \varepsilon_N t,\ |B_{\varepsilon_N t}^{w,N} - w| \le \delta_N\} \\ &\qquad\qquad \times \frac{N}{N'} \left|P_{(1-\varepsilon_N)t}^N(\phi^2)(B_{\varepsilon_N t}^{w,N}) - P_t^N \phi^2(w)\right|\Big) \\ &\quad + \frac{1}{N'} \sum_w \xi_0^N(w) \left|\sum_e p_N(e) \left\{P(\tau^N(w, w+e) > \varepsilon_N t) \frac{N}{N'} - \gamma\right\}\right| P_t^N \phi^2(w). \end{aligned}$$

Recalling the definitions of $X_0^N(\mathbf{1})$, $\kappa_N(t)$ in (4.15) and $\tau^{*,N}$, we may bound the above by

$$\begin{aligned} C_\phi X_0^N(\mathbf{1}) \Big\{ P\left(\left|B_{\varepsilon_N t}^{0,N}\right| > \delta_N\right) \frac{N}{N'} + P(\tau^{*,N} > 2\varepsilon_N t) \frac{N}{N'} \kappa_N(t) \\ + \left|\frac{N}{N'} P(\tau^{*,N} > 2\varepsilon_N t) - \gamma\right| \Big\}. \end{aligned}$$

Use (4.12) and then (K1) and (K2)(b) twice to see the above is at most $X_0^N(\mathbf{1})\rho_N$, where $\rho_N \to 0$ is a sequence of constants independent of $X_0^N$. This establishes (4.11) and completes the proof of Lemma 4.3. □

The second half of the proof of Theorem 4.2 is as follows.

LEMMA 4.4. *Assume* (H1)–(H3). *If* (K1)–(K4) *hold*, *then* (J2) *holds*.

PROOF. To begin to accomplish this we let $R^N(T) = |\{B_s^{0,N}: s \le T\}|$ denote the size of the range of $B_s^{0,N}$ up to time $T$, and derive the following bound.

LEMMA 4.5. *Conditions* (H3)(a) *and* (K4) *imply that* $\sup_N E(R^N(T))/N' < \infty$ *for all* $T > 0$.

PROOF. We start with the observation that

$$ER^N(T) = \sum_x P(B_s^{0,N} = x \text{ for some } s \le T). \tag{4.17}$$

Next, by a last time at $x$ decomposition [see Lemma A.2(ii) in the Appendix],

$$\begin{aligned} &P(B_s^{0,N} = x \text{ for some } s \le T) \\ &\qquad = P(B_T^{0,N} = x) + N \int_0^T P(B_s^{0,N} = x) P(\tau^{*,N} > T - t)\, dt. \end{aligned}$$

Combining these facts gives

$$ER^N(T) = 1 + N \int_0^T P(\tau^{*,N} > t)\, dt.$$

By integrating first over $t \le (N'/N)^2$ and bounding $P(\tau^{*,N} > t)$ by 1 there, and then integrating over $t > (N'/N)^2$, we obtain the estimate

$$\frac{ER^N(T)}{N'} \le \frac{1}{N'} + \frac{N'}{N} + T\frac{N}{N'} P\left(\tau^N > \left(\frac{N'}{N}\right)^2\right). \tag{4.18}$$

By (H3)(a), $N' \le N$, so using (K4) we see that $ER^N(T)/N'$ is bounded. □

PROOF OF (J2.a). If $N' \equiv N$ then $V'_{N,s} \le 1$ and the desired result holds with $g(s) = 1$. Invoking (H2.b), we now suppose that $N'/N \to 0$. Let $\varepsilon_N = (N'/N)^2$. By duality, we have

$$\begin{aligned} E\left[X_s^N(V'_{N,s})\right] &\le \frac{N}{N'^2} \sum_x \sum_y p_N(y - x) P[B_s^{x,N} \in \xi_0^N,\ \tau^N(x, y) > s] \\ &= \frac{N}{N'^2} \sum_w \xi_0^N(w) \sum_e p_N(e) \sum_x P[B_s^{0,N} = w - x,\ \tau^N(0, e) > s] \\ &= X_0^N(\mathbf{1}) \frac{N}{N'} P(\tau^N > 2s). \end{aligned}$$

If $s \ge \varepsilon_N = (N'/N)^2$, (K4) implies this is bounded by $CX_0^N(\mathbf{1})$ for a finite constant $C$. If $0 \le s \le \varepsilon_N$, then the above is at most

$$\frac{N}{N'} X_0^N(\mathbf{1}) \le X_0^N(\mathbf{1}) s^{-1/2}.$$

Therefore (J2)(a) holds with $g(s) = C(1 + s^{-1/2})$. □

PROOF OF (J2)(b). We will warm up by estimating $E[X_s^N(\mathbf{1})^2]$. Recall from Lemma 2.2 that $X_s^N(\mathbf{1})$ is a martingale, with predictable-square function $2\int_0^s X_u^N(V'_{N,u})\,du$. Thus, using (J2)(a), for $s \in [0, T]$,

$$\begin{aligned} EX_s^N(\mathbf{1})^2 &\le X_0^N(\mathbf{1})^2 + 2\int_0^s EX_u^N(V'_N(u))\,du \\ &\le X_0^N(\mathbf{1})^2 + 2X_0^N(\mathbf{1})\int_0^s g(u)\,du \le C_T\Big[X_0^N(\mathbf{1})^2 + 1\Big]. \end{aligned} \tag{4.19}$$

Turning now to the third moments we will again use the particle system duality, but this time for three and four particles. In the proof of (J2.c) we will need to consider four particles, so we pause now for some general definitions. For $A \subset \mathbf{S_N}$ let $\zeta_s^{A,N}$ denote the set of occupied sites in a coalescing random walk system in which: (i) we start with a particle at each site $x \in A$ and (ii) the particles move according to independent random walks $B_t^{x,N}$ until they collide, at which point they stick together for all time. Let

$$\tau^N(A) = \min\{\tau^N(a,b): a, b \in A, a \ne b\}.$$

denote the time of the first collision between two particles in $\zeta_s^{A,N}$. When $A = \{x, y, z\}$, we should write $\tau^N(\{x, y, z\})$ for the collision time, but we will usually drop the braces and write $\tau^N(x, y, z)$ instead.

It follows from the particle system duality that

$$EX_s^N(\mathbf{1})^3 = (N')^{-3}\sum_{x,y,z} P(\zeta_s^{\{x,y,z\},N} \subset \xi_0^N).$$

To bound the right-hand side we start with

$$\begin{aligned} &P\Big[\zeta_s^{\{x,y,z\},N} \subset \xi_0^N, \tau^N(x,y,z) > s\Big] \\ &\quad = P\Big[\{B_s^{x,N}, B_s^{y,N}, B_s^{z,N}\} \subset \xi_0^N, \tau^N(x,y,z) > s\Big] \\ &\quad \le P(B_s^{x,N} \in \xi_0^N)P(B_s^{y,N} \in \xi_0^N)P(B_s^{z,N} \in \xi_0^N). \end{aligned}$$

It is immediate from (a) of Lemma 2.3 that

$$(N')^{-3}\sum_{x,y,z} P(B_s^{x,N} \in \xi_0^N)P(B_s^{y,N} \in \xi_0^N)P(B_s^{z,N} \in \xi_0^N) = X_0^N(\mathbf{1})^3. \tag{4.20}$$

By decomposing the event $\{\tau^N(x, y, z) \le s\}$ according to which pair of walks collides first, then using symmetry and translation, we see that

$$\begin{aligned} &\sum_{x,y,z} P\big(\zeta_s^{\{x,y,z\},N} \subset \xi_0^N, \tau^N(x,y,z) \le s\big) \\ &\quad = 3\sum_{x,y',z'} P\big(\zeta_s^{\{0,y',z'\},N} \subset \xi_0^N - x, \tau^N(0,y') = \tau^N(0,y',z') \le s\big) \end{aligned} \tag{4.21}$$

where $y' = y - x$, $z' = z - x$, and $0, y'$ correspond to the pair of walks which collides first. Ignoring the restriction that the walk starting from $z$ avoids the other walks up to the coalescing time, using the strong Markov property, and dropping the primes from $y$ and $z$, we obtain

$$\sum_{x,\,y,\,z} P(\zeta_s^{\{0,\,y,\,z\},\,N} \subset \xi_0^N - x,\, \tau^N(0,y) = \tau^N(0,y,z) \le s)$$
$$\le \sum_{x,\,y,\,z}\sum_{w,\,w'} \int_0^s P(\tau^N(0,y) \in du,\, B_u^{0,\,N} = w)$$
$$\times P(B_u^{z,\,N} = w')P(\zeta_{s-u}^{\{w,\,w'\},\,N} \subset \xi_0^N - x).$$

Do the sum over $z$ and rearrange the other sums to see the above equals

$$\text{(4.22)} \qquad \sum_{y,\,w} \int_0^s P(\tau^N(0,y) \in du,\, B_u^{0,\,N} = w) \sum_{x,\,w'} P(\zeta_{s-u}^{\{w,w'\},N} \subset \xi_0^N - x).$$

Changing variables and doing the sums with the help of (a) of Lemma 2.3 and (4.19) gives

$$(N')^{-2} \sum_{x,\,w'} P(\zeta_{s-u}^{\{w,\,w'\},\,N} \subset \xi_0^N - x) = N'^{-2} \sum_{x',\,w''} P(\zeta_{s-u}^{\{x',\,w''\},\,N} \subset \xi_0^N)$$
$$= EX_{s-u}^N(\mathbf{1})^2 \le C_T(X_0^N(\mathbf{1})^2 + 1).$$

With the dependence on $u$ eliminated we can sum to get

$$(N')^{-1} \sum_{y,\,w} \int_0^s P(\tau^N(0,y) \in du,\, B_u^{0,\,N} = w)$$
$$= (N')^{-1} \sum_y P(\tau^N(0,y) \le s) = (N')^{-1} ER^N(2s),$$

the 2 coming from the fact that $B_t^{0,\,N} - B_t^{y,\,N}$ makes jumps at rate $2N$. Plugging the last two equations into (4.22) and using Lemma 4.5, we see that

$$\text{(4.23)} \qquad \frac{1}{(N')^3} \sum_{x,\,y,\,z} P(\zeta^{x,\,y,\,z,\,N} \subset \xi_0^N,\, \tau^N(x,y,z) \le s) \le C_T(X_0^N(\mathbf{1})^2 + 1).$$

The estimates (4.20) and (4.23) imply (J2)(b). □

PROOF OF (J2)(c). We need to estimate $E[X_s^N(V'_{N,\,s})X_s^N(\mathbf{1})^2]$. If $N' \equiv N$ then $V'_{N,\,s} \le 1$ so this follows from (J2)(b). So again we invoke (H2)(b) and suppose that $N'/N \to 0$. Let $\varepsilon_N = (N'/N)^2$. For $s \le \varepsilon_N$, (J2)(b) implies this expectation is at most

$$\text{(4.24)} \qquad \frac{N}{N'} E(X_s^N(\mathbf{1})^3) \le \frac{N}{N'} C(X_0^N(\mathbf{1})^3 + 1) \le C(X_0^N(\mathbf{1})^3 + 1)s^{-1/2}$$

for some constant $C$ independent of $N$.

Having disposed of $s \le \varepsilon_N$ we can and will for the rest of the proof suppose $s > \varepsilon_N$. Using duality, we have

$$E(X_s^N(V'_{N,s})X_s^N(\mathbf{1})^2) \le \frac{N}{(N')^4} \sum_{x,y,z,x'} p_N(x-x')P(\zeta_s^{\{x,y,z\},N} \subset \xi_0^N, \tau^N(x,x') > s). \tag{4.25}$$

An upper bound for this is $\Sigma_1^N(s) + \Sigma_2^N(s) + \Sigma_3^N(s)$, where

$$\Sigma_1^N(s) = \frac{N}{(N')^4} \sum_{x,y,z} P(\zeta_s^{\{x,y,z\},N} \subset \xi_0^N, \tau^N(x,y,z) \le \varepsilon_N),$$

$$\Sigma_2^N(s) = \frac{N}{(N')^4} \sum_{x,y,z,x'} p_N(x-x')P[\zeta_s^{\{x,y,z\},N} \subset \xi_0^N, \tau^N(x',y,z) \le \varepsilon_N < \tau^N(x,y,z)],$$

$$\Sigma_3^N(s) = \frac{N}{(N')^4} \sum_{x,y,z,x'} p_N(x-x')P[\zeta_s^{\{x,y,z\},N} \subset \xi_0^N, \tau^N(x',x,y,z) > \varepsilon_N],$$

and in the last we have used $\tau_N(x,x') > s > \varepsilon_N$.

I. To handle $\Sigma_1^N(s)$, note that by symmetry and translation it equals

$$3\frac{N}{(N')^4} \sum_{x,y,z} P(\zeta_s^{\{x,y,z\},N} \subset \xi_0^N, \tau^N(x,z) = \tau^N(x,y,z) \le \varepsilon_N)$$
$$= 3\frac{N}{(N')^4} \sum_{x,y,z} P(\zeta_s^{\{0,y,z\},N} \subset \xi_0^N - x, \tau^N(0,z) = \tau^N(0,y,z) \le \varepsilon_N).$$

Decompose according to the value of $\tau^N(0,z) = u$, and the positions $B_u^{0,N}$ and $B_u^{y,N}$, to obtain

$$\Sigma_1^N(s) \le 3\,\frac{N}{(N')^4} \sum_{x,y,z} \sum_{w,w'} \int_0^{\varepsilon_N} P(\tau^N(0,z) \in du, B_u^{0,N} = w) \times P(B_u^{y,N} = w')P(\zeta_{s-u}^{\{w,w'\},N} \subset \xi_0^N - x).$$

Doing the sum over $y$, rearranging and changing variables, we see the right side above equals

$$3\frac{N}{(N')^2} \sum_{w,z} \int_0^{\varepsilon_N} P(\tau^N(0,z) \in du, B_u^{0,N} = w)\frac{1}{(N')^2} \times \sum_{x',w''} P\big(\zeta_{s-u}^{\{x',w''\},N} \subset \xi_0^N\big). \tag{4.26}$$

Using duality and (4.19), then summing over $w$ and doing the integral over $u$, we transform the above (for $s \le T$) into

$$3\frac{N}{(N')^2} \sum_{w,z} \int_0^{\varepsilon_N} P(\tau^N(0,z) \in du, B_u^{0,N} = w)EX_{s-u}^N(1)^2$$

$$\leq \frac{N}{(N')^2} \sum_z P(\tau^N(0,z) \leq \varepsilon_N) C_T(X_0^N(\mathbf{1})^2 + 1)$$

$$= \frac{N}{(N')^2} E[R^N(2\varepsilon_N)] C_T(X_0^N(\mathbf{1})^2 + 1).$$

Clearly $R^N(2\varepsilon_N)$ is bounded by the number of jumps of $B^{0,N}$ up to time $2\varepsilon_N$, that is, a Poisson variable with mean $2\varepsilon_N N$. Thus, for an appropriate constant $C_T$, we have for $T \geq s > \varepsilon_N$,

$$\Sigma_1^N(s) \leq C_T(X_0^N(\mathbf{1})^2 + 1)\left(\frac{N^2}{(N')^2}\right)\varepsilon_N = C_T(X_0^N(\mathbf{1})^2 + 1), \tag{4.27}$$

since $\varepsilon_N = (N'/N)^2$.

II. For $\Sigma_2^N(s)$, note that on the event $\{\tau^N(x', y, z) \leq \varepsilon_N < \tau^N(x, y, z)\}$, $y$ and $z$ cannot be the first pair to hit. Therefore, on this event,

$$\tau^N(\{x', y, z\}) = \tau^N(x', y) \wedge \tau^N(x', z).$$

Using symmetry, $\Sigma_2^N(s)$ is at most

$$\frac{2N}{(N')^4} \sum_{x,y,z,x'} p_N(x - x') P\Big[\zeta_s^{\{x,y,z\},N} \subset \xi_0^N,$$
$$\tau^N(x', y) = \tau^N(x', y, z) \leq \varepsilon_N < \tau^N(x, y, z)\Big].$$

Changing variables $e = x - x'$, $y' = y - x'$, $z' = z - x'$, and weakening various inequalities, we see the above is bounded by

$$\frac{2N}{(N')^4} \sum_{x',y',z',e} p_N(e) P\Big[\zeta_s^{\{e,y',z'\},N} \subset \xi_0^N - x',$$
$$\tau^N(0, y') = \tau^N(0, y', z') \leq \varepsilon_N < \tau^N(y', z')\Big]$$
$$\leq \frac{2N}{(N')^4} \sum_{x',y',z'} P\Big[\zeta_s^{\{y',z'\},N} \subset \xi_0^N - x',\ \tau^N(0, y') = \tau^N(0, y', z') \leq \varepsilon_N\Big].$$

Dropping the primes from the variables $x'$, $y'$, $z'$ and introducing $u$, $w$ and $w'$ to account for the possible values of $\tau^N(0, y)$, $B_u^{0,N}$ and $B_u^{z,N}$, the strong Markov property implies that the above is most

$$\frac{2N}{(N')^4} \sum_{x,y,z,w,w'} \int_0^{\varepsilon_N} P(\tau^N(0, y) \in du,\ B_u^{0,N} = w)$$
$$\times P(B_u^{z,N} = w') P[\zeta_{s-u}^{\{w,w'\},N} \subset \xi_0^N - x].$$

If we sum over $z$ to get rid of $P(B_u^{z,N} = w')$, then we get the same expression as in (4.26), and so we have

$$\Sigma_2^N(s) \leq C_T(X_0^N(\mathbf{1})^2 + 1). \tag{4.28}$$

III. Finally, we come to the estimation of $\Sigma_3^N(s)$ for $s > \varepsilon_N$. Translating by $x$, setting $e = x' - x$ and computing as we did for the two previous sums, we obtain

$$\begin{aligned}
\Sigma_3^N(s) &= \frac{N}{(N')^4} \sum_{e, x, y, z} p_N(-e) P(\zeta_s^{\{0, y, z\}, N} \subset \xi_0^N - x, \tau^N(e, 0, y, z) > \varepsilon_N) \\
&\le \frac{N}{(N')^4} \sum_{e, x, y, z} \sum_{w, w', w''} p_N(e) P(B_{\varepsilon_N}^{0, N} = w, \tau^N(0, e) > \varepsilon_N) \\
&\qquad \times P(B_{\varepsilon_N}^{y, N} = w') \\
&\qquad \times P(B_{\varepsilon_N}^{z, N} = w'') P[\zeta_{s-\varepsilon_N}^{\{w, w', w''\}, N} \subset \xi_0^N - x] \\
&= \frac{N}{(N')^4} \sum_{w, e} p_N(e) P(B_{\varepsilon_N}^{0, N} = w, \tau^N(0, e) > \varepsilon_N) \\
&\qquad \times \sum_{x, w', w''} P(\zeta_{s-\varepsilon_N}^{\{w, w', w''\}, N} \subset \xi_0^N - x).
\end{aligned}$$

A change of variables gives

$$\begin{aligned}
&\frac{1}{(N')^3} \sum_{x, w', w''} P(\zeta_{s-\varepsilon_N}^{\{w, w', w''\}, N} \subset \xi_0^N - x) \\
&\qquad = \frac{1}{(N')^3} \sum_{a, b, c} P(\zeta_{s-\varepsilon_N}^{\{a, b, c\}, N} \subset \xi_0^N) = EX_{s-\varepsilon_N}^N(\mathbf{1})^3.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\Sigma_3^N(s) &\le \frac{N}{N'} EX_{s-\varepsilon_N}^N(\mathbf{1})^3 \sum_{w, e} p_N(e) P(B_{\varepsilon_N}^{0, N} = w, \tau^N(0, e) > \varepsilon_N) \\
&= \frac{N}{N'} EX_{s-\varepsilon_N}^N(\mathbf{1})^3 \sum_{e} p_N(e) P(\tau^N(0, e) > \varepsilon_N) \\
&\le C_s(X_0^N(\mathbf{1})^3 + 1) \frac{N}{N'} P(\tau^{*, N} > 2\varepsilon_N)
\end{aligned}$$

by the third moment estimate (J2)(b). Recalling $\varepsilon_N = (N'/N)^2$ and using (K4), there is a constant $C_T$ such that if $T \ge s > \varepsilon_N$,

$$\Sigma_3^N(s) \le C_T(X_0^N(\mathbf{1})^3 + 1). \tag{4.29}$$

Combining (4.25) with (4.27)–(4.29) gives (J2)(c) with $g(s) = C_s(1 + s^{-1/2})$. Note that the boundedness and integrability conditions on $g$ are clear as we may take $C_s$ to be increasing in $s$. □

For the proof of Theorem 4.2, combine Lemmas 4.3 and 4.4.

For the proof of Theorem 1.4, combine Theorems 3.5, 4.1 and 4.2.

**5. Proofs of Theorems 1.1–1.3.** As (H1)–(H3) hold in all of these settings, by Theorem 1.4 it suffices to show that the conditions (K1)–(K4) of Section 1 hold for the parameter combinations $(p_N, M_N, N', \gamma)$ being considered. That is, it suffices to prove the following result.

THEOREM 5.1. *Conditions* (K1)–(K4) *hold in each of the following cases*:

(a) *The long range kernels given by* (M1), *with* $M_N$ *satisfying* (1.5), $N' \equiv N$ *and* $\gamma = 1$.

(b) *The fixed kernels in* (M2), *with* $M_N \equiv 1$, *and* $N'$, $\gamma$ *given by* (1.6) *and* (1.7).

(c) *The long range kernels of* (M1) *in* $d = 2$, *with* $M_N$, $N'$ *and* $\gamma$ *given in* (1.8)–(1.10).

PROOF. We will take the three parts of Theorem 5.1 in order. To prove (a), we will first obtain an estimate on the spread out random walk kernels given in (M1) which will be used to verify (K3).

Recall that $\mathbf{S_N} = \mathbf{Z}^d/(M_N\sqrt{N})$.

LEMMA 5.2. *Assume* $W_N$ *is uniformly distributed over* $(\mathbf{Z}^d/M_N)\cap I$, *where* $I = [-1, 1]^d \setminus \{0\}$. *There is a* $C$ *such that for all* $t \geq 0$,

$$\sup_{x\in\mathbf{S_N}} P(B_t^{0,N} = x) \leq \exp\left(\frac{-Nt}{2}\right) + \frac{C}{M_N^d(Nt+1)^{d/2}}.$$

PROOF. By Lemma 2.4 of Bramson, Durrett and Swindle (1989) (and a trivial rescaling), there is a constant $C$ such that for all $N \in \mathbf{N}$ and $u > 0$,

$$\sup_x P\left(B_u^{0,N} \in x + \left[\frac{-1}{\sqrt{N}}, \frac{1}{\sqrt{N}}\right]^d\right) \leq \frac{C}{(Nu+1)^{d/2}}. \tag{5.1}$$

Let $T_1$ be the first jump time of $B_u^{0,N}$. Then, by the Markov property,

$$\begin{aligned} P(B_t^{0,N} = x) &= P(B_t^{0,N} = x,\ T_1 > t/2) + P(B_t^{0,N} = x, T_1 \leq t/2) \\ &\leq \exp\left(\frac{-Nt}{2}\right) + \int_0^{t/2} P(T_1 \in du) \sum_e p_N(e) P(B_{t-u}^{e,N} = x). \end{aligned}$$

The probability of $B_{t-u}^{e,N} = x$ is the same as that of $B_{t-u}^{0,N} - x = -e$. Furthermore,

$$\sum_e p_N(e) P(B_{t-u}^{0,N} - x = -e) = \frac{1}{(2M_N+1)^d - 1} P\left(B_{t-u}^{0,N} - x \in \left[\frac{-1}{\sqrt{N}}, \frac{1}{\sqrt{N}}\right]^d\right).$$

Using (5.1), we now have

$$P(B_t^{0,N} = x) \leq \exp\left(\frac{-Nt}{2}\right) + \frac{1}{(2M_N+1)^d - 1} \cdot \frac{C}{(Nt/2+1)^{d/2}}. \qquad \square$$

PROOF OF THEOREM 5.1 (a). Recall that $N \equiv N'$, which makes (K4) trivial. By considering separately the three cases, $d = 1$, $d = 2$ and $d \geq 3$, it is clear that (K3) follows from Lemma 5.2. To check (K2) we begin by observing that for any $\varepsilon_N > 0$, a simple estimate using Chebyshev's inequality gives

$$P(|B^{0,N}_{t\varepsilon_N}| > \delta) \leq \frac{E|B^{0,N}_{t\varepsilon_N}|^2}{\delta^2} \leq \frac{Ct\varepsilon_N}{\delta^2}.$$

This shows that (K2.a) holds for any $\varepsilon_N \to 0$. For (K2)(b), we will show that for any fixed $t > 0$,

$$P(\tau^{*,N} \leq t) \to 0 \quad \text{as } N \to \infty. \tag{5.2}$$

This will imply that (K1) holds with $\gamma = 1$ and (K2)(b) holds for any $\varepsilon_N \leq 1$.

For any $\delta_N > 0$, the number of jumps the walk $B^{*,N}_t$ takes up to time $\delta_N$ is Poisson distributed with mean $N\delta_N$. The probability it hits 0 on any given jump is at most $1/M_N^d$, so the probability that it hits 0 in its first $k$ jumps is at most $k/M_N^d$. Therefore,

$$P(\tau^{*,N} \leq \delta_N) \leq \exp(-N\delta_N) \sum_{k=1}^{\infty} \frac{(N\delta_N)^k}{k!} \frac{k}{M_N^d} = \frac{N\delta_N}{M_N^d}. \tag{5.3}$$

On the other hand, a last time at 0 decomposition [Lemma A.2(iii)] gives

$$P(\tau^{*,N} \in (\delta_N, t]) \leq P(B^{*,N}_t = 0) + N \int_{\delta_N}^t P(B^{*,N}_u = 0)\, du.$$

By Lemma 5.2, $P(B^{*,N}_t = 0) \to 0$ as $N \to \infty$. Since $\int_{\delta_N}^t N \exp(-Nu/2)\, du \leq 2\exp(-N\delta_N/2)$, Lemma 5.2 also implies that

$$\begin{aligned} &N \int_{\delta_N}^t P(B^{*,N}_u = 0)\, du \\ &\leq 2\exp\left(\frac{-N\delta_N}{2}\right) + \begin{cases} \dfrac{C}{M_N}(Nt+1)^{1/2}, & \text{in } d = 1, \\ \dfrac{C}{M_N^2} \log\Big((Nt+1)/(N\delta_N+1)\Big), & \text{in } d = 2, \\ \dfrac{C}{M_N^d}(1/(N\delta_N+1)^{(d/2)-1}), & \text{in } d \geq 3. \end{cases} \end{aligned} \tag{5.4}$$

If we choose $\delta_N$ such that $N\delta_N \to \infty$ as $N \to \infty$, the first term on the right side of (5.4) will tend to 0. It is a simple matter to check, using the assumptions on $M_N$, that the second term on the right side of (5.4) and the right side of (5.3) tend to 0 if we set $\delta_N = N^{-1/2}$ in $d = 1$, $\delta_N = \log N/N$ in $d = 2$ and $\delta_N = M_N/N$ in $d = 3$. This shows that (5.2) holds and completes the proof of (a). □

PROOF OF THEOREM 5.1(b). We consider now the fixed kernel models in $d \geq 2$ that have $M_N \equiv 1$.

*Case* 1. In $d \geq 3$, $N \equiv N'$, so (K4) is trivial. To begin to check (K3), let $Z_t^x$ be a rate-one continuous time random walk on $\mathbf{Z}^d$ with step distribution $p(x)$, and let $\tau_0(x)$ denote the first hitting time of 0 for $Z_t^x$. By a standard local limit theorem [see (A.7) in the Appendix], there is a finite $C$ such that for all $N$,

$$P(B_t^{0,N} = 0) = P(Z_{tN}^0 = 0) \leq \frac{C}{(Nt+1)^{d/2}}. \tag{5.5}$$

Condition (K3) clearly follows from (5.5) and the simple bound (see Lemma A.3)

$$P\Big(B_t^{0,N} = x\Big) \leq P\Big(B_t^{0,N} = 0\Big). \tag{5.6}$$

For (K2)(a), we note that for any $\varepsilon_N > 0$, a simple estimate using Chebyshev's inequality gives

$$P\Big(\Big|B_{t\varepsilon_N}^{0,N}\Big| > \delta\Big) \leq \frac{E\Big|B_{t\varepsilon_N}^{0,N}\Big|^2}{\delta^2} \leq \frac{Ct\varepsilon_N}{\delta^2}. \tag{5.7}$$

This shows that (K2)(a) holds for any $\varepsilon_N \to 0$. To check the last two conditions, we note that

$$P\Big(\tau^{*,N} > t\Big) \to \gamma_e \tag{5.8}$$

and that if $\varepsilon_N N \to \infty$, then

$$P\Big(\tau^{*,N} > \varepsilon_N t\Big) = \sum_{x \in \mathbf{Z}^d} p(x) P(\tau_0(x) > \varepsilon_N N t) \to \gamma_e. \tag{5.9}$$

Thus, choosing $\varepsilon_N \to 0$ so that $\varepsilon_N N \to \infty$, (K1) holds with $\gamma = \gamma_e$ and (K2)(b) is satisfied. □

*Case* 2. In $d = 2$, we are assuming that $M_N \equiv 1$ and $N = N' \log N$. Let $Z_t^x$ and $\tau_0(x)$ be as above. Write $a(t) \sim b(t)$ as $t \to \infty$ if and only if $\lim_{t\to\infty} a(t)/b(t) = 1$. By a standard local limit theorem [see Lemma A.3(i) below],

$$P\Big(Z_t^0 = 0\Big) \sim \frac{1}{2\pi\sigma^2 t} \qquad \text{as } t \to \infty. \tag{5.10}$$

Furthermore, by Lemma A.3(ii),

$$\sum_{x \in \mathbf{Z}^d} p(x) P(\tau_0(x) > t) \sim \frac{2\pi\sigma^2}{\log t} \qquad \text{as } t \to \infty. \tag{5.11}$$

To see that (K3) holds, we use (5.10), obtaining

$$\frac{N}{\log N} P\Big(B_t^{0,N} = 0\Big) = \frac{N}{\log N} P\Big(Z_{tN}^0 = 0\Big) \leq \frac{C}{\log N} \to 0$$

as $N \to \infty$. Using (5.6) now (which remains valid for $d = 2$) gives (K3).

To check that (K1) holds with $\gamma = 2\pi\sigma^2$, we use (5.11) to conclude

$$(\log N)P(\tau^{*,N} > t) \sim 2\pi\sigma^2 \frac{\log N}{\log(Nt)} \sim 2\pi\sigma^2.$$

For (K2)(a), we can again use the Chebyshev estimate in (5.7), obtaining

$$(\log N)P\Big(\Big|B^{0,N}_{\varepsilon_N t}\Big| > \delta\Big) \le (\log N)\frac{Ct\varepsilon_N}{\delta^2} \to 0$$

provided $(\log N)\varepsilon_N \to 0$. For (K2)(b) we note that if $\log(\varepsilon_N)/\log N \to 0$, then another use of (5.11) gives

$$(\log N)P(\tau^{*,N} > \varepsilon_N t) \sim 2\pi\sigma^2 \frac{\log N}{\log N + \log \varepsilon_N + \log t} \sim 2\pi\sigma^2.$$

If we let $\varepsilon_N = 1/(\log N)^2$, then both parts of (C2) hold. Finally, (K4) is a trivial consequence of (5.11). □

PROOF OF THEOREM 5.1(c). Here we consider the long-range kernels of (M1) for $d = 2$, assuming now that $M_N^2/(\log N) \to \rho \in [0, \infty)$ as $N \to \infty$. We consider first Case 1.

*Case* 1. Suppose $\rho > 0$, where $N' \equiv N$ and $\gamma = 1/\{1+(3/2\pi\rho)\}$. Since $N' \equiv N$, condition (K4) is trivial, and condition (K3) follows easily from Lemma 5.2. Reasoning as in (5.7) we see that (K2)(a) holds for any $\varepsilon_N \to 0$.

The last two conditions require some additional information not provided by Lemma 5.2. To introduce this, let $G_N(t) = \int_0^t P(B_s^{0,N} = 0)\,ds$. By Lemma A.4 of the Appendix, if $\varepsilon_N \to 0$ with $\varepsilon_N M_N^2 \to \infty$, then for any fixed $T > 0$ and $t_N \in [\varepsilon_N T, T]$,

$$\lim_{N\to\infty} NG_N(t_N) = 1 + \frac{3}{2\pi\rho}. \tag{5.12}$$

Let $H_N(t) = P(\tau^{*,N} > t)$. We claim that

$$\lim_{N\to\infty} NG_N(t_N)H_N(t_N) = 1. \tag{5.13}$$

Given this, (K1) and (K2)(b) follow immediately from (5.12).

To prove (5.13), we start with the decomposition [see Lemma A.2(i)]

$$1 = P(B_t^{0,N} = 0) + N\int_0^t P(B_s^{0,N} = 0)H_N(t-s)\,ds. \tag{5.14}$$

Since $H_N(t)$ is nonincreasing in $t$, it is clear that

$$NG_N(t)H_N(t) \le 1. \tag{5.15}$$

Therefore, to prove (5.13) it suffices to prove

$$\liminf_{N\to\infty} NG_N(t_N)H_N(t_N) \ge 1. \tag{5.16}$$

The decomposition in (5.14) implies that

$$1 \le P(B_{2t}^{0,N} = 0) + NH_N(t)\int_0^t P(B_s^{0,N} = 0)\,ds + N\int_t^{2t} P(B_s^{0,N} = 0)\,ds.$$

Rearranging this, and setting $t = t_N$, we obtain

$$NG_N(t_N)H_N(t_N) \ge 1 - P(B_{2t_N}^{0,N} = 0) - N[G_N(2t_N) - G_N(t_N)]. \tag{5.17}$$

It is easy to see from Lemma 5.2 that if $\varepsilon_N N \to \infty$ (which is weaker than the already assumed condition, $\varepsilon_N M_N^2 \to \infty$) then

$$P(B_{2t_N}^{0,N} = 0) \to 0 \tag{5.18}$$

as $N \to \infty$, and also that

$$N[G_N(2t_N) - G_N(t_N)] \le 2\exp\left(\frac{-N\varepsilon_N T}{2}\right) + \frac{C\log 2}{M_N^2} \to 0. \tag{5.19}$$

Combining (5.17)–(5.19) gives (5.16).

*Case* 2. Suppose $M_N \to \infty$ and $M_N^2/(\log N) \to \rho = 0$ as $N \to \infty$. Let $N' = NM_N^2/\log N$, and $\gamma = 2\pi/3$. As before, Lemma 5.2 implies that (K3) holds, and the Chebyshev estimate in (5.7) gives

$$\frac{N}{N'}P\left(\left|B_{\varepsilon_N t}^{0,N}\right| > \delta\right) \le \frac{N}{N'}\frac{Ct\varepsilon_N}{\delta^2},$$

which certainly tends to 0 as $N \to \infty$ if we set

$$\varepsilon_N = \left(\frac{N'}{N}\right)^2 = \left(\frac{M_N^2}{\log N}\right)^2.$$

Therefore, (K2)(a) holds for this choice of $\varepsilon_N$.

Lemma A.4 implies that, for any fixed $T > 0$ and $\varepsilon_N'$ satisfying $\varepsilon_N' \to 0$ and $\varepsilon_N' M_N^2 \to \infty$ as $N \to \infty$, for $t_N \in [\varepsilon_N' T, T]$,

$$\lim_{N\to\infty} \frac{M_N^2}{\log N} NG_N(t_N) = \frac{3}{2\pi}. \tag{5.20}$$

We claim that (5.20) in fact holds for $t_N \in [\varepsilon_N T, T]$. This is because Lemma 5.2 implies that

$$N[G_N(T) - G_N(\varepsilon_N T)] \le 2\exp\left(\frac{-N\varepsilon_N T}{2}\right) + \frac{C}{M_N^2}\log(1/\varepsilon_N),$$

and this estimate shows that

$$\frac{M_N^2}{\log N} N[G_N(T) - G_N(\varepsilon_N T)] \to 0 \quad \text{as } N \to \infty.$$

Thus, (5.20) holds for $t_N \in [\varepsilon_N T, T]$.

Conditions (K1) and (K2)(b) now follow from (5.20) and (5.13) (which only required $N\varepsilon_N \to \infty$ and so still hold for the current choice of $\varepsilon_N$). Finally, (K4) follows from (K2)(b) since $\varepsilon_N = (N'/N)^2$. □

## APPENDIX

We collect here several "well-known" results, which we have referred to, but that seem difficult to document in the literature. First, we show in Theorem A.1 that the martingale problem $(\mathrm{MP})_{X_0}^{b,\sigma^2}$ for super-Brownian motion using smooth test functions of compact support is well posed. Next, we state and prove extensions to the continuous time setting of several discrete time random walk facts. These include some simple last exit time formulas (see Lemma A.2) and asymptotic hitting time formulas (see Lemma A.3). Finally, we derive asymptotics for the two-dimensional long-range random walk kernels in the "critical" case $M_N^2/\log N \to \rho \in [0,\infty)$ (Lemma A.4).

**A1.** Existence and uniqueness in law of solutions to $(\mathrm{MP})_{X_0}^{b,\sigma^2}$ clearly reduces to existence and uniqueness of a law on the canonical space $\Omega_{X,C}$ under which the coordinate maps $X_t(\omega) = \omega(t)$ satisfy $(\mathrm{MP})_{X_0}^{b,\sigma^2}$ with respect to the canonical right-continuous filtration $\mathscr{F}_t$ containing the null sets of the law in question.

It is at times useful to deal with smooth test functions of compact support in the martingale problem calculations. Although the following result is certainly "well known" most standard references seem to use slightly larger classes of test functions. For example, Dawson (1993) (Theorem 6.1.3, Example 7.1.3) uses the domain of the strong generator of Brownian motion in the space of continuous functions on $\mathbf{R}^d$ with limits at the point at infinity.

THEOREM A.1. *For each $X_0$ in $M_F(\mathbf{R}^d)$, $b > 0$, and $\sigma^2 > 0$, there is a unique probability $P_{X_0}^{b,\sigma^2}$ on $\Omega_{X,C}$ under which $X$ satisfies* $(\mathrm{MP})_{X_0}^{b,\sigma^2}$.

PROOF. Existence is immediate from Example 7.1.3 of Dawson (1993) which uses the larger class of test functions described above. For uniqueness we have to show that our smaller class is good enough. Let $P$ be a probability on $\Omega_{X,C}$ under which $X$ satisfies the above martingale problem. If $\{h_n\}$ is the sequence of functions considered in Lemma 3.3, let $g_n = \mathbf{1} - h_n \uparrow \mathbf{1}$ with $g_n \in C_0^\infty(\mathbf{R}^d)$. If $T_k = \inf\{t: X_t(\mathbf{1}) \geq k\}$, with $T_k = \infty$ if there is no such $t$, then $T_k \uparrow \infty$ $P$-a.s. By $(\mathrm{MP})_{X_0}^{b,\sigma^2}$ and monotone convergence,

$$E(X_{t\wedge T_k}(\mathbf{1})) = \lim_{n\to\infty} E(X_{t\wedge T_k}(g_n)) = \lim_{n\to\infty} X_0(g_n) + E\biggl(\int_0^{t\wedge T_k} X_s\biggl(\frac{\sigma^2}{2}\Delta g_n\biggr)\,ds\biggr).$$

Since $\|\Delta g_n\|_\infty$ is uniformly bounded and $\Delta g_n$ converges pointwise to 0, we may use dominated convergence (note that $X_{s\wedge T_k}(|\sigma^2 \Delta g_n|) \leq ck$) to see that the above gives

$$E(X_{t\wedge T_k}(\mathbf{1})) = \lim_{n\to\infty} X_0(g_n) = X_0(\mathbf{1}).$$

Now use Fatou's lemma to conclude that

$$E(X_t(\mathbf{1})) \le X_0(\mathbf{1}) < \infty. \tag{A.1}$$

Therefore, for $T > 0$, $(\mathrm{MP})_{X_0}^{b,\sigma^2}$ and Doob's maximal inequality show that

$$E\Big(\sup_{t\le T}(M_t(g_m) - M_t(g_n))^2\Big) \le cE\Big(\int_0^T X_s(b(g_m - g_n)^2)\,ds\Big) \to 0 \tag{A.2}$$

as $m, n \to \infty$, by (A.1) and dominated convergence. Note (by dominated convergence again) that for all $t \ge 0$ a.s.,

$$M_t(g_n) = X_t(g_n) - X_0(g_n) - \int_0^t X_s\Big(\frac{\sigma^2}{2}\Delta g_n\Big)\,ds \to X_t(\mathbf{1}) - X_0(\mathbf{1}) \equiv M_t(\mathbf{1})$$

as $n \to \infty$. By (A.2), the above convergence is uniform in $t \le T$ in $L^2$, and so $M_t(\mathbf{1})$ is a continuous, square integrable $(\mathscr{F}_t)$-martingale satisfying for all $t \ge 0$,

$$\langle M(\mathbf{1})\rangle_t = \lim_{n\to\infty}\langle M(g_n)\rangle_t = \int_0^t X_s(b\mathbf{1})\,ds,$$

where the convergence is in $L^2$. By continuity this establishes $(\mathrm{MP})_{X_0}^{b,\sigma^2}$ for $\phi$ in $C_0^\infty(\mathbf{R}^d) \cup \{\mathbf{1}\}$. Let $D_0$ denote the linear span of this set of functions. It follows from (A.2) that $\langle M(g_n), M(\phi)\rangle_t$ converges to $\langle M(\mathbf{1}), M(\phi)\rangle_t$ in $L^2$, and from this one may easily establish $(\mathrm{MP})_{X_0}^{b,\sigma^2}$ for $\phi \in D_0$. $D_0$ is a core for $A$, the generator of the Brownian semigroup (with variance $\sigma^2$) on the space of continuous functions with limits at $\infty$ [see Proposition 5.1.1 of Ethier and Kurtz (1986)]. It follows that $(\mathrm{MP})_{X_0}^{b,\sigma^2}$ holds for $\phi$ in $D(A)$, the domain of $A$ [use the square integrability of $X_t(\mathbf{1}) = X_0(\mathbf{1}) + M_t(\mathbf{1})$ established above]. This is enough to conclude that $P$ is the law of super-Brownian motion with branching rate $b$ and diffusion rate $\sigma^2$ by standard results which are easy to find [e.g., by Theorem 6.1.3 of Dawson (1993) and Itô's lemma]. □

**A2.** We temporarily adopt the notation of Chung, (1967). Let $x(t)$, $t \ge 0$ be a rate $r$ continuous time random walk on $\mathbf{Z}^d$ with step distribution $(p_i)$, $i \in \mathbf{Z}^d$, with $p_0 = 0$. That is, we let $p_{ij} = p_{j-i}$ and let $x(t)$ have probability transition semigroup $p_{ij}(\cdot)$, where

$$p_{ij}(t) = \sum_{n=0}^{\infty} e^{-rt}(rt)^n p_{ij}^{(n)}/n!, \qquad i, j \in \mathbf{Z}^d.$$

We define the random variables $\tau_i = \inf\{t: x(t) = i\}$, $\gamma_i(t) = \sup\{s \le t: x(s) = i\}$, where $\inf\varnothing = \sup\varnothing = \infty$. That is, $\tau_i$ is the first hitting time of $i$, and $\gamma_i(t)$ is the last time at $i$ before time $t$. We also define $H(t) = \sum_i p_i P^i(\tau_0 > t)$ and the "taboo" probabilities ${}_ip_{kj}(t) = P^k(x(t) = j, \tau_i > t)$. Our next step is to prove the following "last exit time" formulas.

LEMMA A.2. (i) $1 = p_{ii}(t) + r\int_0^t p_{ii}(s)H(t-s)\,ds, \quad i \in \mathbf{Z}^d.$

(ii) $$P^i(\tau_j \le t) = p_{ij}(t) + r\int_0^t p_{ij}(s)H(t-s)\,ds, \qquad i, j \in \mathbf{Z}^d,\ i \ne j.$$

(iii) $$P^i(\tau_j \in (\delta, t]) \le p_{ij}(t) + r\int_\delta^t p_{ij}(s)H(t-s)\,ds, \qquad i, j \in \mathbf{Z}^d,\ i \ne j,\ 0 \le \delta \le t.$$

PROOF. By Theorem II.12.3 of Chung (1967), $p_{ij}(t) = \int_0^t p_{ii}(s)g_{ij}(t-s)\,ds$, for $i \ne j$, where

$$g_{ij}(s) = \lim_{\delta\to 0}\frac{1}{\delta}\sum_{k\ne i} p_{ik}(\delta)\,_i p_{kj}(t).$$

It follows easily in our setting from the bounded convergence theorem that

$$g_{ij}(t) = r\sum_{k\ne i} p(k-i)_i\, p_{kj}(t). \tag{A.3}$$

Therefore,

$$p_{ij}(t) = r\int_0^t p_{ii}(s)\sum_{k\ne i} p(k-i)_i\, p_{kj}(t-s)\,ds, \qquad i \ne j,$$

and summation over $j$ yields $1 = p_{ii}(t) + r\int_0^t p_{ii}(s)H(t-s)\,ds$. This proves (i).

For (ii), we start with the simple decomposition

$$P^i(\tau_j \le t) = p_{ij}(t) + P^i(\tau_j < t, x(t) \ne j) = p_{ij}(t) + P^i(\gamma_j(t) < t). \tag{A.4}$$

By the Markov property, consideration of the first hitting time of $j$ gives

$$P^i(\gamma_j(t) < t) = \int_0^t P^i(\tau_j \in ds)P^j(\gamma_j(t-s) < t-s).$$

By (i) above, and the fact that $P^j(\gamma_j(t) < t) = 1 - p_{jj}(t)$, we have

$$P^j(\gamma_j(t) < t) = r\int_0^t p_{jj}(t-s)H(s)\,ds.$$

Therefore,

$$\begin{aligned} P^i(\gamma_j(t) < t) &= r\int_0^t P^i(\tau_j \in ds)\int_0^{t-s} p_{jj}(t-s-u)H(u)\,du \\ &= r\int_0^t p_{ij}(t-u)H(u)\,du. \end{aligned} \tag{A.5}$$

In the last line we have used the strong Markov property at $\tau_j$. By combining (A.4) and (A.5) we obtain (ii). To prove (iii), take differences in (ii) and use the fact that $H$ is nonincreasing. □

**A3.** Here we give a brief sketch of how to adapt standard asymptotics for discrete time random walk to the continuous time setting. Let $p(x, y) = p(x-y)$ be an irreducible, symmetric random walk kernel on $\mathbf{Z}^d$, such that $p(0) = 0$ and $\sum_{x\in\mathbf{Z}^d} x^i x^j p(x) = \delta_{ij}\sigma^2$. Let $B_t$ denote rate-one continuous time random walk with step distribution $p(x)$, and let $\tau = \tau_0$, so that

$$H(t) = \sum_x p(x) P^x(\tau > t).$$

Recall that $a(t) \sim b(t)$ as $t \to \infty$ means that $\lim_{t\to\infty} a(t)/b(t) = 1$.

LEMMA A.3. *For every $t > 0$ and $x \in \mathbf{Z}^d$, $p_t(0, x) \le p_t(0, 0)$. As $t \to \infty$,*

(i) $p_t(0,0) \sim \left(\dfrac{1}{2\pi\sigma^2 t}\right)^{d/2}$,

(ii) $H(t) \sim \dfrac{2\pi\sigma^2}{\log t}$ *in $d = 2$.*

PROOF. Let $\phi(\theta) = \sum_x p(x)\exp(-ix\cdot\theta)$ be the characteristic function of the step distribution $p(\cdot)$. Then

$$p_t(0, x) = (2\pi)^{-d} \int_{[-\pi,\pi]^d} \exp(-ix\cdot\theta)\exp(-t(1-\phi(\theta)))\,d\theta.$$

This representation and the fact that $\phi(\theta)$ is real (by the symmetry of $p$) imply that $p_t(0, x) \le p_t(0, 0)$. By a standard local limit theorem [see P9 on page 79 of Spitzer (1976)],

$$p_n(0,0) \sim \left(\frac{1}{2\pi\sigma^2 n}\right)^{d/2} \quad \text{as } n \to \infty \text{ with } n \in \mathbf{N}. \tag{A.6}$$

By the Fourier representation for $p_t(0,0)$ it is apparent that $p_t(0,0)$ is nonincreasing in $t$. This fact and (A6) easily imply (i). It also follows that there is a finite constant $C$ such that

$$p_t(0,0) \le C/(1 + t^{d/2}), \qquad t \ge 0. \tag{A.7}$$

Let $G(t) = \int_0^t p_s(0,0)\,ds$, and observe by (i) that $G(t) \sim (\log t)/2\pi\sigma^2$ as $t \to \infty$ in $d = 2$. Given this, in order to prove (ii), it suffices to prove $G(t)H(t) \to 1$ as $t \to \infty$. To do so we employ the last exit time decomposition of Lemma A.2,

$$1 = p_t(0,0) + \int_0^t p_s(0,0) H(t-s)\,ds. \tag{A.8}$$

Since $H(s)$ is nonincreasing, it follows from (A.8) that

$$1 \ge G(t)H(t). \tag{A.9}$$

To obtain a bound in the other direction, we employ (A.8) again, and obtain, for any $a > 1$,

$$1 = p_{at}(0,0) + \int_0^{at} p_s(0,0) H(at-s)\,ds$$

$$\le p_{at}(0,0) + \int_0^{(a-1)t} p_s(0,0)H(t)\,ds + \int_{(a-1)t}^{at} p_s(0,0)\,ds$$
$$= p_{at}(0,0) + G((a-1)t)H(t) + G(at) - G((a-1)t).$$

Thus, for any $a > 1$, $H(t)G((a-1)t) \ge 1 - p_{at}(0,0) - (G(at) - G((a-1)t)$. By the estimate (A.7),

$$\limsup_{t\to\infty}[G(at) - G((a-1)t)] \le C\log(a/(a-1))$$

and $\lim_{t\to\infty} G((a-1)t)/G(t) = 1$. Since $p_{at}(0,0) \to 0$ as $t \to \infty$, it follows that

$$\liminf_{t\to\infty} G(t)H(t) \ge 1 - C\log(a/(a-1)), \qquad a > 1.$$

Letting $a \uparrow \infty$, we obtain $\liminf_{t\to\infty} G(t)H(t) \ge 1$. This and (A.9) complete the proof of (ii). □

**A4.** The next result gives the asymptotics needed to verify that the $d = 2$ long-range random kernels with $M_N^2 = O(\log N)$ satisfy the conditions (K1) and (K2). Let $\Lambda_M = [-M, M]^d \cap (\mathbf{Z}^d \setminus \{0\})$, and let $U_t^{x,M}$ denote continuous time rate-one random walk with initial state $x$ and step distribution that is uniform over $\Lambda_M$.

LEMMA A.4. *Assume $d = 2$ and $M_N^2/(\log N) \to \rho \in [0,\infty)$ as $N \to \infty$. Let $\varepsilon_N > 0$, and assume that $\varepsilon_N \to 0$ and $\varepsilon_N M_N^2 \to \infty$ as $N \to \infty$. Then for fixed $T > 0$, for any $t_N \in [\varepsilon_N TN, TN]$,*

$$\lim_{N\to\infty} \frac{M_N^2}{\log N}\int_0^{t_N} P(U_s^{0,M_N} = 0)\,ds = \rho\biggl(1 + \frac{3}{2\pi\rho}\biggr).$$

PROOF. To simplify notation, we will write $M$ for $M_N$. The characteristic function of $U_t^{0,M}$ is $\exp(-t(1-\varphi_M))$, where

$$\varphi_M(\theta) = \frac{1}{|\Lambda_M|}\sum_{x\in\Lambda_M}\exp(ix\cdot\theta) = \frac{1}{|\Lambda_M|}\Biggl[\prod_{j=1}^{2}\frac{\sin((M+\frac12)\theta_j)}{\sin(\frac12\theta_j)} - 1\Biggr].$$

By the inversion formula [see P6.3 in Spitzer (1976)],

$$P(U_s^{0,M} = 0) = (2\pi)^{-2}\int_{B(\pi)}\exp(-s(1-\varphi_M(\theta)))\,d\theta,$$

where $B(r) = [-r,r]^2$. It follows easily that

$$\text{(A.10)}\quad \int_0^{t_N} P(U_s^{0,M} = 0)\,ds = (2\pi)^{-2}\int_{B(\pi)}\frac{1-\exp(-t_N(1-\varphi_M(\theta)))}{1-\varphi_M(\theta)}\,d\theta.$$

We will evaluate this integral by splitting $B(\pi)$ into several regions, and estimating the integrand over these regions. To do this, we will rely on the following properties of $\varphi_M$, which can be easily derived using the explicit form of $\varphi_M$.

(P1) For all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\frac{1-\varphi_M(\theta)}{M^2|\theta|^2/6} \in (1-\varepsilon, 1+\varepsilon) \qquad \text{for all } \theta \in B(\delta/M).$$

(P2) For all $\delta > 0$ there exists $a = a(\delta) > 0$ and finite $C = C(\delta) < 1$ such that

$$|\varphi_M(\theta)| \le C \qquad \text{for all } \theta \in B(a) \setminus B(\delta/M).$$

(P3) For all $\varepsilon' > 0$ there exists a finite $C = C'_\varepsilon$ such that

$$|\varphi_M(\theta)| \le \frac{C}{M^2} \qquad \text{for all } \theta \in B(\pi) \setminus B(\varepsilon').$$

Let us now fix $0 < \varepsilon < 1$. To simplify notation, we define

$$\psi_M(\theta) = \frac{1-\exp(-t_N(1-\varphi_M(\theta)))}{1-\varphi_M(\theta)},$$

and also the annular region, $A(r,s) = B(s) \setminus B(r)$ for $0 < r < s$. We note that $\psi_M$ is real and positive. By a simple inequality, $\psi_M(\theta) \le t_N$ everywhere, and since $t_N \le TN$,

$$\text{(A.11)} \qquad \frac{1}{4\pi^2}\int_{B(\varepsilon/\sqrt{N})} \psi_M(\theta)\,d\theta \le T\,\frac{\varepsilon^2}{\pi^2}.$$

Now choose $\delta$ as in (P1). Then, for $\theta \in A(\varepsilon/\sqrt{N}, \delta/M)$, since $t_N \ge \varepsilon_N TN$,

$$t_N(1-\varphi_M(\theta)) \ge t_N(1-\varepsilon)M^2|\theta|^2/6 \ge \frac{\varepsilon^2(1-\varepsilon)T}{6}\,\varepsilon_N M^2,$$

which tends to infinity as $N \to \infty$ by assumption. For large $N$, $\exp[-\varepsilon^2(1-\varepsilon)T\delta_N M^2/6] \le \varepsilon$. Since (P1) also implies that for $\theta \in B(\delta/M)$,

$$1-\varphi_M(\theta) \le (1+\varepsilon)M^2|\theta|^2/6,$$

we have that, for large $N$, and $\theta \in A(\varepsilon/\sqrt{N}, \delta/M)$,

$$\psi_M(\theta) \ge \frac{1-\varepsilon}{1+\varepsilon}\,\frac{6}{M^2|\theta|^2}.$$

Therefore,

$$\begin{aligned}
\frac{1}{4\pi^2}\int_{A(\varepsilon/\sqrt{N},\,\delta/M)} \psi_M(\theta)\,d\theta &\ge \frac{1-\varepsilon}{1+\varepsilon}\,\frac{3}{2\pi^2 M^2}\int \frac{1\{2\varepsilon/\sqrt{N} \le |\theta| \le \delta/2M\}}{|\theta|^2}\,d\theta \\
&\ge (1-2\varepsilon)\,\frac{3}{\pi M^2}\int_{2\varepsilon/\sqrt{N}}^{\delta/2M} \frac{dr}{r} \\
&= (1-2\varepsilon)\,\frac{3}{\pi M^2}\,\log\!\left(\frac{\delta\sqrt{N}}{4\varepsilon M}\right).
\end{aligned}$$

Using the fact that $M \to \infty$ and $M^2/\log N \to \rho \in [0, \infty)$ as $N \to \infty$, it follows that

$$\text{(A.12)} \qquad \liminf_{N\to\infty} \frac{M^2}{\log N} \frac{1}{4\pi^2} \int_{A(\varepsilon/\sqrt{N},\, \delta/M)} \psi_M(\theta)\, d\theta \geq (1-2\varepsilon)\frac{3}{2\pi}.$$

A corresponding analysis shows that

$$\text{(A.13)} \qquad \limsup_{N\to\infty} \frac{M^2}{\log N} \frac{1}{4\pi^2} \int_{A(\varepsilon/\sqrt{N},\, \delta/M)} \psi_M(\theta)\, d\theta \leq (1+2\varepsilon)\frac{3}{2\pi}.$$

Now choose $a$ as in (P2). Then $\psi \leq 1/(1-C(\delta))$ on $A(\delta/M, a)$. We may now choose $\varepsilon' > 0$, $\varepsilon' < \min(a, \varepsilon)$ small enough so that

$$\text{(A.14)} \qquad \frac{1}{4\pi^2} \int_{A(\delta/M,\, \varepsilon')} \psi(\theta)\, d\theta \leq \frac{(\varepsilon')^2}{\pi^2} \frac{1}{1-C(\delta)} < \varepsilon.$$

It follows easily from (P3) that $\psi_M \to 1$ uniformly on $A(\varepsilon', \pi)$, and therefore

$$\text{(A.15)} \qquad \lim_{N\to\infty} \frac{1}{4\pi^2} \int_{A(\varepsilon',\, \pi)} \psi_M(\theta)\, d\theta = 1 - \frac{(\varepsilon')^2}{\pi^2}.$$

If we now combine the estimates (A.11), (A.14) and (A.15), recalling that $\varepsilon' < \varepsilon$, we have

$$\text{(A.16)} \quad \limsup_{N\to\infty} \left| \frac{1}{4\pi^2} \int_{B(\pi)\setminus A(\varepsilon\sqrt{N},\, \delta/M)} \psi_M(\theta)\, d\theta - 1 \right| \leq \varepsilon + (T+1)\varepsilon^2/\pi^2.$$

Clearly, (A.12), (A.13) and (A.16) give the desired result. □

## REFERENCES


BRAMSON, M., COX, J. T. and LE GALL, J.-F. (1999). Super-Brownian limits of voter model clusters. Unpublished manuscript.

BRAMSON, M., DURRETT, R. and SWINDLE, G. (1989). Statistical mechanics of crabgrass. *Ann. Probab.* **17** 441–481.

BRAMSON, M. and GRIFFEATH, D. (1980).

BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42.

CHUNG, K. L. (1967). *Markov Chains with Stationary Transition Probabilities.* Springer, New York.

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

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

DAWSON, D. A. (1992) Infinitely divisible random measures and superprocesses. In *Stochastic Analysis and Related Topics* 1–129. Birkhäuser, Boston.

DAWSON, D. A. (1993). Measure–valued Markov processes. *Ecole d'été de Probabilités de Saint Flour* XXI. *Lecture Notes in Math.* **1541** 1–260. Springer, New York.

DAWSON, D. A. and HOCHBERG, K. J. (1979). The carrying dimension of a stochastic measure diffusion. *Ann. Probab.* **7** 693–703.

DAWSON, D. A. and PERKINS, E. (1991). Historical processes. *Mem. Amer. Math. Soc.* **454**.

DAWSON, D. A. and PERKINS, E. (1998). Measure-valued processes and renormalization of branching particle systems. In *Stochastic Partial Differential Equations: Six Perspectives* (R. Carmona and B. Rozovskii, eds.) 45–106 Amer. Math. Soc., Providence, RI.

DERBEZ, E. and SLADE, G. (1997). Lattice trees and super-Brownian motion. *Canad. Math. Bull.* **40** 19–38.

DERBEZ, E., VAN DER HOFSTAD, R. and SLADE, G. (1998). Unpublished manuscript.

DURRETT, R. (1995). Ten lectures on particle systems. *Ecole d'été de Probabilités de Saint Flour XXIII. Lecture Notes in Math.* **1608** 97–201. Springer, New York.

DURRETT, R. and PERKINS, E. (1999). Rescaled contact processes converge to super-Brownian motion for $d \geq 2$. *Probab. Theory Related Fields.* To appear.

DYNKIN, E. B. (1994). *An Introduction to Branching Measure-Valued Processes.* Amer. Math. Soc., Providence.

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

FITZSIMMONS, P. J. (1988). Construction and regularity of measure-valued Markov branching processes. *Israel J. Math.* **64** 337–361.

GRIFFEATH, D. (1978). *Additive and Cancellative Interacting Particle Systems. Lecture Notes in Math.* **724**. Springer, Berlin.

JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes.* Springer, Berlin.

JAKUBOWSKI, A. (1986). On the Skorohod topology. *Ann. Inst. H. Poincaré Math. Statist. Probab.* **22** 263–285.

KONNO, N. and SHIGA, T. (1988). Stochastic differential equations for some measure-valued diffusions. *Probab. Theory Related Fields* **79** 201–225.

KURTZ, T. and PROTTER, P. (1996). Weak convergence of stochastic integrals and differential equations. Infinite-dimensional case. *Probabilistic Models for Nonlinear Partial Differential Equations. Lecture Notes in Math.* **1627** 197–285. Springer, New York.

LE GALL, J.-F. (1994). Mouvement brownien, processus du branchement et superprocessus. Prepublication 247 du Laboratoire de Probabilités.

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

MUELLER, C. and TRIBE, R. (1995). Stochastic PDE's arising from the long range contact process and long range voter model. *Probab. Theory Related Fields* **102** 519–546.

PRESUTTI, E. and SPOHN, H. (1983). Hydrodynamics of the voter model. *Ann. Probab.* **11** 867–875.

REIMERS, M. (1989). One dimensional stochastic partial differential equations and the branching measure valued diffusions. *Probab. Theory Related Fields* **81** 319–340.

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

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

J. T. COX
DEPARTMENT OF MATHEMATICS
SYRACUSE UNIVERSITY
SYRACUSE NEW YORK 13244
E-mail: mathdept@suvm.acs.syr.edu

R. DURRETT
DEPARTMENT OF MATHEMATICS
MALOTT HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

E. A. PERKINS
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
1984 MATHEMATICS ROAD
VANCOUVER, BRITISH COLUMBIA V6T 1Z2
CANADA