Convergence of branching particle systems and sparse voter models to super-Brownian motion

Jan M. Swart

November 22, 2010

Abstract

These are lecture notes for a sequence of lectures given at the *Seminar on stochastic* evolution equations held October-December 2010 at the Institute of Information Theory and Automation of the ASCR (ÚTIA) in Prague. They are based on a series of papers by Cox, Durrett, Merle and Perkins on the convergence of rescaled sparse voter models to super-Brownian motion.

Contents

1	A bit of Markov process theory	2
	1.1 Generators and the martingale problem	2
	1.2 Covariances and quadratic variation	2
	1.3 Continuous martingales and Itô's formula	3
2	Branching particle systems and super-Brownian motion	4
	2.1 Branching particle systems	4
	2.2 Super Brownian motion	5
	2.3 Scaling limits of branching particle systems	6
	2.4 The log-Laplace semigroup	8
3	Convergence of sparse voter models	8
	3.1 Voter models	8
	3.2 Convergence in recurrent dimensions	9
	3.3 Convergence in dimension two	13
4	Convergence of Neuhauser-Pacala models	14
	4.1 Neuhauser-Pacala models	14
	4.2 Convergence in transient dimensions	15
	4.3 Convergence in dimension two	18

1 A bit of Markov process theory

1.1 Generators and the martingale problem

Let $X = (X_t)_{t\geq 0}$ be a Markov process with finite state space S that jumps from x to y with rate r(x, y). Let $P_t(x, y)$ denote the transition probabilities of X, i.e., $P_t(x, y)$ is the probability that $X_t = y$ given that we start the process in $X_0 = x$. For any function $f: S \to \mathbb{R}$, we define

$$P_t f(x) := \sum_{y \in S} P_t(x, y) f(y).$$
(1.1)

Then $(P_t)_{t\geq 0}$ is a semigroup of linear operators acting on the finite dimensional space of real functions on S. We have $P_t = e^{tG}$, where

$$Gf(x) := \sum_{y \in S} r(x, y) \big(f(y) - f(x) \big).$$
(1.2)

is the generator of X. In particular, $P_t = 1 + tG + O(t^2)$ so

$$\mathbb{E}^{x}[f(X_{t})] = P_{t}f(x) = f(x) + tGf(x) + O(t^{2}) \quad \text{as } t \to 0,$$
(1.3)

where \mathbb{E}^x denotes expectation with respect to the law of the process started from $X_0 = x$.

It is not hard to prove that if X is a Markov process with jump rates r(x, y) (and arbitrary initial law), then for each $f: S \to \mathbb{R}$, the process

$$M_t^f := f(X_t) - \int_0^t \mathrm{d}s \, Gf(X_s) \tag{1.4}$$

is a martingale with respect to the filtration generated by $(X_t)_{t\geq 0}$. Conversely, any S-valued stochastic process with this property is a Markov process with jump rates r(x, y). Formula (1.4) is described in words by saying that the process X solves the martingale problem for the operator G.

1.2 Covariances and quadratic variation

Recall that if $M = (M_t)_{t\geq 0}$ and $N = (N_t)_{t\geq 0}$ are square integrable martingales with rightcontinuous sample paths, then there exists a unique predictable, nondecreasing process $\langle M, N \rangle$ with right-continuous sample paths, starting at zero, such that

$$M_t N_t - \langle M, N \rangle_t \tag{1.5}$$

is a martingale. We call $\langle M, N \rangle$ the predictable covariation process of M and N. In particular, $\langle M \rangle := \langle M, M \rangle$ is called the predictable quadratic variation process of M. One has

$$4\langle M,N\rangle = \langle M+N,M+N\rangle - \langle M-N,M-N\rangle.$$
(1.6)

If X is a Markov process as in the pervious section and M^f, M^g are martingales as in (1.4), then

$$\langle M^f, M^g \rangle_t = \int_0^t \mathrm{d}s \, \Gamma(f, g)(X_s), \tag{1.7}$$

where

$$\Gamma(f,g) = G(fg) - fGg - gGf = \sum_{y \in S} r(x,y) \big(f(y) - f(x) \big) \big(g(y) - g(x) \big).$$
(1.8)

The function $\Gamma(f,g)$ here also appears in the well-known covariance formula:

$$\operatorname{Cov}(f(X_t), g(X_t)) = \operatorname{Cov}(P_t f(X_0), P_t f(X_0)) + \int_0^t \mathrm{d}s \,\mathbb{E}[\Gamma(P_s f, P_s g)(X_{t-s})].$$
(1.9)

1.3 Continuous martingales and Itô's formula

We say that a process V is of bounded variation if

$$\sup \sum_{k=1}^{n} |V_{t_k} - V_{t_{k-1}}| < \infty \qquad (t > 0), \tag{1.10}$$

where the supremum is over all partitions $0 = t_0 < \cdots < t_n = t$ of [0, t]. By definition, a *semimartingale* is a process $X = (X_t)_{t \ge 0}$ that can be written as

$$X_t = M_t + V_t \qquad (t \ge 0),$$
 (1.11)

where M is a right-continuous square integrable martingale and V is a continuous adapted process of bounded variation, starting at zero. Since continuous adapted processes are predictable, such a decomposition is unique. For any two *continuous* semimartingales X, Y, one has

$$\langle X, Y \rangle_t = \lim_{n \to \infty} \sum_{k=1}^n (X_{t_k^n} - X_{t_{k-1}^n}) (Y_{t_k^n} - Y_{t_{k-1}^n}) \qquad (t \ge 0).$$
 (1.12)

where the limit is in probability and $0 = t_0^n < \cdots < t_n^n = t$ is a sequence of increasingly fine partitions of [0, t], such that $\sup_{k=1}^n |t_k^n - t_{k-1}^n| \to 0$.¹ If X is a continuous semimartingale and Y is a continuous process of bounded variation, then $\langle X, Y \rangle_t = 0$ $(t \ge 0)$. In particular, one has

$$\langle X, Y \rangle_t = \langle M, N \rangle_t \quad \text{when} \quad X_t = M_t + V_t, \quad Y_t = N_t + W_t$$
(1.13)

and V, W are continuous process of bounded variation.

Local semimartingales are defined similarly to semimartingales, in the usual way. Stochastic integrals w.r.t. (continuous) semimartingales are well-defined and yield again (continuous) semimartingales. If $\vec{X} = X^1, \ldots, X^n$ are continuous local semimartingales and $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, then $f(\vec{X})$ is again a continuous local semimartingale. In fact, *Itô's formula* says that

$$f\left(\vec{X}_{t}\right) = f\left(\vec{X}_{0}\right) + \sum_{i=1}^{n} \int_{0}^{t} \left(\frac{\partial}{\partial x^{i}}f\right)(\vec{X}_{s}) \,\mathrm{d}X_{s}^{i} + \frac{1}{2} \sum_{i,j=1}^{n} \int_{0}^{t} \left(\frac{\partial^{2}}{\partial x^{i}\partial x^{j}}f\right)(\vec{X}_{s}) \,\mathrm{d}\langle X^{i}, X^{j}\rangle_{s}.$$
 (1.14)

¹If X and Y are square integrable martingales but their sample paths are not continuous, then the limit in (1.12) also exists, is denoted by $[X, Y]_t$, and is called the *covariation process* of X and Y. Also in this case, it is true that $X_tY_t - [X, Y]_t$ is a martingale, but in general $[X, Y] \neq \langle X, Y \rangle$ and [X, Y] need not be predictable.

2 Branching particle systems and super-Brownian motion

2.1 Branching particle systems

Let $\mathcal{N}(\mathbb{Z}^d)$ be the space of finite counting measure on \mathbb{Z}^d ; equivalently, we may view an element of $\mathcal{N}(\mathbb{Z}^d)$ as a function $\rho : \mathbb{Z}^d \to \mathbb{N}$ such that $|\rho| := \sum_{x \in \mathbb{Z}^d} \rho(x) < \infty$. Let $p : \mathbb{Z}^d \to [0, 1]$ be a probability distribution on \mathbb{Z}^d satisfying

(i)
$$p(x) = p(-x),$$

(ii) $\sum_{x} x_i x_j p(x) = \sigma^2 \delta_{ij},$
(iii) $\sum_{x} |x|^3 p(x) < \infty$
(iv) $\{x : p(x) > 0\}$ generates $\mathbb{Z}^d.$
(2.1)

Let $(\rho_t)_{t\geq 0}$ be a Markov process with values in $\mathcal{N}(\mathbb{Z}^d)$ (and cadlag sample paths) defined by the following jump rates:

$$\begin{array}{ll}
\rho \mapsto \rho - \delta_x + \delta_y & \text{with rate} & p(y - x)\rho(x), \\
\rho \mapsto \rho + \delta_x & \text{with rate} & b\rho(x), \\
\rho \mapsto \rho - \delta_x & \text{with rate} & d'\rho(x),
\end{array}$$
(2.2)

where $b, d' \ge 0$ are constants. If we interpret $\rho_t(x)$ as the number of particles at site x at time t, then this says that particles jump from x to y with rate p(y - x), particles branch into two new particles with rate b, and particles die with rate d'.

Note from this description of the process that there is no interaction between the particles. In particular, it is easy to see that our process has the *branching property:* if $(\rho'_t)_{t\geq 0}$ and $(\rho''_t)_{t\geq 0}$ are independent realizations of the process, started in initial states ρ'_0, ρ''_0 , then

$$\rho_t := \rho'_t + \rho''_t \qquad (t \ge 0) \tag{2.3}$$

is a Markov process with the same dynamics as ρ'_t and ρ''_t , started in $\rho'_0 + \rho''_0$.

Branching particle systems are 'linear' interacting particle systems, in the following sense. Let $(S_t)_{t\geq 0}$ be the semigroup of a branching particle system on \mathbb{Z}^d with jump kernel p, branching rate b and death rate d'. Let $(P_t)_{t\geq 0}$ be the semigroup of a random walk on \mathbb{Z}^d that jumps from x to y with rate p(y - x). For each bounded $\phi : \mathbb{Z}^d \to \mathbb{R}$, define a 'linear' function $l_{\phi} : \mathcal{N}(\mathbb{Z}^d) \to \mathbb{R}$ by

$$l_{\phi}(\rho) := \sum_{x \in \mathbb{Z}^d} \rho(x)\phi(x).$$
(2.4)

Then S_t maps linear functions into linear functions, and in fact:

$$S_t l_\phi = l_{e^{(b-d')t} P_t \phi} \qquad (t \ge 0).$$
 (2.5)

In view of this, the random walk that jumps from x to y with rate p(y - x) is called the *underlying motion* of the branching particle system.

Setting $\phi = 1$ in (2.5) show that the expected total mass $|\rho_t| := \sum_x \rho_t(x)$ satisfies

$$\mathbb{E}[|\rho_t|] = e^{(b-d')t} \mathbb{E}[|\rho_0|] \qquad (t \ge 0).$$
(2.6)

In fact, the total mass process $(|\rho_t|)_{t\geq 0}$ is an autonomous Markov proces with values in \mathbb{N} , that jumps

$$\begin{array}{ll}
n \mapsto n+1 & \text{with rate} & b, \\
n \mapsto n-1 & \text{with rate} & d'.
\end{array}$$
(2.7)

The process is called *subcritical*, *critical* or *supercritical* depending on whether b < d', b = d', or b > d'. It turns out that the subcritical and critical processes, started in a finite initial state, die out almost surely, while the supercritical process has a positive probability to survive.

Branching particle system can also be defined for infinite initial states. In this case, the long-time behavior of the critical process is more subtle. It turns out that in dimensions 1 and 2 the process still dies out but in dimensions 3 and more there exist nontrivial invariant laws.

2.2 Super Brownian motion

Let $\mathcal{M}(\mathbb{R}^d)$ be the space of finite measures on \mathbb{R}^d , equipped with the topology of weak convergence. For $\mu \in \mathcal{M}(\mathbb{R}^d)$ and f a measurable real function on \mathbb{R}^d , we adopt the notation

$$\mu(f) := \int \mu(\mathrm{d}x) f(x). \tag{2.8}$$

Recall that the Laplacian Δ is the differential operator defined by

$$\Delta\phi(x) := \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} \phi(x) \qquad (x \in \mathbb{R}^d).$$
(2.9)

By definition, super-Brownian motion is the unique $\mathcal{M}(\mathbb{R}^d)$ -valued Markov process $(X_t)_{t\geq 0}$ with continuous sample paths such that for each $\phi \in \mathcal{C}^3_{\mathrm{b}}(\mathbb{R}^d)$ (the space of three times continuously differentiable functions on $\phi : \mathbb{R}^d \to \mathbb{R}$ such that ϕ and its derivatives up to third order are bounded), the process

$$M_t(\phi) := X_t(\phi) - \int_0^t \mathrm{d}s \, X_s\left(\frac{1}{2}\sigma^2 \Delta \phi + \beta \phi\right) \tag{2.10}$$

is a continuous, square integrable martingale with quadratic variation process

$$\langle M(\phi) \rangle_t = \int_0^t \mathrm{d}s \, X_s(\gamma \phi^2).$$
 (2.11)

We call the nonnegative contants σ^2 the diffusion constant, β the growth rate, and γ the activity (or branching rate) of the process.

Super Brownian motion is very much like a continuum analogue of a branching particle system and in fact occurs a the scaling limits of such sytems, as we will see below. Many of the elementary properties of super-Brownian motion are analogous to those of branching particle systems. In particular, super-Brownian motion has the branching property, it is a linear system (with a Brownian underlying motion) and its total mass process $(|X_t|)_{t\geq 0}$ is an autonomous Feller diffusion, i.e., it is equal in law to the unique solution $(F_t)_{t\geq 0}$ of the stochastic differential equation

$$\mathrm{d}F_t = \beta F_t \,\mathrm{d}t + \sqrt{\gamma F_t} \,\mathrm{d}B_t. \tag{2.12}$$

Super Brownian motion is called subcritical, critical or supercritical depending on whether $\beta < 0, \beta = 0$ or $\beta > 0$. Its long-time behavior is similar to that of branching particle systems.

To give the reader a feeling why (2.10) and (2.11) might be enough to characterize a Markov process, we note that if $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, then by Itô's

formula (1.14):

$$f(X_t(\phi_1), \dots, X_t(\phi_n)) = f(X_0(\phi_1), \dots, X_0(\phi_n)))$$

+ $\sum_{i=1}^n \int_0^t (\frac{\partial}{\partial x_i} f)(X_s(\phi_1), \dots, X_s(\phi_n)) dX_s(\phi_i)$
+ $\frac{1}{2} \sum_{i,j=1}^n \int_0^t (\frac{\partial^2}{\partial x_i \partial x_j} f)(X_s(\phi_1), \dots, X_s(\phi_n)) d\langle X(\phi_i), X(\phi_j) \rangle_s,$ (2.13)

where by (1.6) and (2.11),

$$\langle X(\phi_i), X(\phi_j) \rangle_t = \int_0^t \mathrm{d}s \, X_s(\gamma \phi_i \phi_j).$$
(2.14)

Using moreover (2.10) we see that

$$f(X_t(\phi_1), \dots, X_t(\phi_n)) = f(X_0(\phi_1), \dots, X_0(\phi_n)))$$

+ $\sum_{i=1}^n \int_0^t (\frac{\partial}{\partial x_i} f)(X_s(\phi_1), \dots, X_s(\phi_n)) X_s(\frac{1}{2}\Delta\phi_i + \beta\phi_i) ds$
+ $\frac{1}{2} \sum_{i,j=1}^n \int_0^t (\frac{\partial^2}{\partial x_i \partial x_j} f)(X_s(\phi_1), \dots, X_s(\phi_n)) X_s(\gamma\phi_i\phi_j) ds$
+ stochastic integral w.r.t. $M_s(\phi_i)$. (2.15)

It follows that as $t \to 0$,

$$\mathbb{E}^{\rho} \Big[f \Big(X_t(\phi_1), \dots, X_t(\phi_n) \Big) \Big] = f \Big(\rho(\phi_1), \dots, \rho(\phi_n) \Big) \Big) + t \sum_{i=1}^n (\frac{\partial}{\partial x_i} f) \Big(\rho(\phi_1), \dots, \rho(\phi_n) \Big) \rho \Big(\frac{1}{2} \sigma^2 \Delta \phi_i + \beta \phi_i \Big) + \frac{1}{2} t \sum_{i,j=1}^n (\frac{\partial^2}{\partial x_i \partial x_j} f) \Big(\rho(\phi_1), \dots, \rho(\phi_n) \Big) \rho(\gamma \phi_i \phi_j) + O(t^2).$$

$$(2.16)$$

By (1.3), this means that from (2.10) and (2.11) we can deduce the action of the generator of super-Brownian motion on smooth functions of the form $\rho \mapsto f(\rho(\phi_1), \ldots, \rho(\phi_n))$.

2.3 Scaling limits of branching particle systems

Theorem 1 (Scaling limit of branching particle systems) Let ε_n, δ_n be positive constants, converging to zero. Let $(\rho_t^n)_{t\geq 0}$ be branching particle systems on \mathbb{Z}^d with jump kernel p, branching rate b_n , and death rate d'_n . Assume that

$$\varepsilon_n^{-2}(b_n - d'_n) \underset{n \to \infty}{\longrightarrow} \beta \quad and \quad \delta_n \varepsilon_n^{-2}(b_n + d'_n) \underset{n \to \infty}{\longrightarrow} \gamma.$$
 (2.17)

Define $\mathcal{M}(\mathbb{R}^d)$ -valued processes $(X_t^n)_{t\geq 0}$ by

$$X_{\varepsilon_n^2 t}^n := \delta_n \sum_{x \in \mathbb{Z}^d} \rho_t^n(x) \delta_{\varepsilon_n x} \qquad (t \ge 0).$$
(2.18)

Assume that $X_0^n \Rightarrow X_0$ for some $X_0 \in \mathcal{M}(\mathbb{R}^d)$ and let $(X_t)_{t\geq 0}$ be the super-Brownian motion with this initial state and diffusion constant σ^2 , growth rate β and activity γ . Then

$$\mathbb{P}[(X_t^n)_{t\geq 0}\in\cdot]\underset{n\to\infty}{\Longrightarrow}\mathbb{P}[(X_t)_{t\geq 0}\in\cdot].$$
(2.19)

Remarks By (2.17), we have $(b_n - d'_n)/(b_n + d'_n) \sim \delta_n \to 0$ as $n \to \infty$, so we in the limit we are looking at near-critical branching. For given $\beta \neq 0$, $\gamma > 0$, and b_n, d'_n , the scaling factors ε_n, δ_n are up to asymptotic equivalence uniquely determined by (2.17). If $\beta = 0$, then we have one free scaling parameter in our choice of ε_n, δ_n . This is reflected in the fact that for any a > 0, critical super-Brownian motion is invariant under a simultaneous scaling of space by a, time by a^2 , and mass by a^2 .

Sketch of the proof For $\phi \in \mathcal{C}^3_{\mathrm{b}}(\mathbb{R}^d)$ and $\varepsilon > 0$, we define a linear function $f^{\varepsilon,\delta}_{\phi} : \mathcal{N}(\mathbb{Z}^d) \to \mathbb{R}$ by

$$f_{\phi}^{\varepsilon,\delta}(\rho) := \delta \sum_{x \in \mathbb{Z}^d} \rho(x) \phi(\varepsilon x).$$
(2.20)

Let G be the generator of a branching particle system on \mathbb{Z}^d with jump kernel p, branching rate b, and death rate d' and let $\Gamma(f,g)$ be defined as in (1.8). Then $\varepsilon^{-2}G$ is the generator of the process with time speeded up by a factor ε^{-2} and

$$\varepsilon^{-2}Gf_{\phi}^{\varepsilon,\delta}(\rho) = \delta\varepsilon^{-2}\sum_{x,y} p(y-x)\rho(x)\big(\phi(\varepsilon y) - \phi(\varepsilon x)\big) + \delta\varepsilon^{-2}(b-d')\sum_{x}\rho(x)\phi(\varepsilon x)$$

$$= \delta\sum_{x}\rho(x)\big(\frac{1}{2}\Delta_{\varepsilon}\phi(\varepsilon x) + \varepsilon^{-2}(b-d')\phi(\varepsilon x)\big).$$
(2.21)

Speeding up time by a factor ε^{-2} means that we also must multiply $\Gamma(f,g)$ from (1.8) by a factor ε^{-2} . Then

$$\varepsilon^{-2}\Gamma(f_{\phi}^{\varepsilon,\delta}, f_{\phi}^{\varepsilon,\delta})(\rho) = \delta^{2}\varepsilon^{-2}\sum_{x,y} p(y-x)\rho(x) \big(\phi(\varepsilon y) - \phi(\varepsilon x)\big)^{2} + \delta^{2}\varepsilon^{-2}(b+d')\sum_{x} \rho(x) \big(\phi(\varepsilon x)\big)^{2} = \delta\sum_{x} \rho(x) \big(\delta R_{\varepsilon}(\phi)(\varepsilon x) + \delta\varepsilon^{-2}(b+d')(\phi(\varepsilon x))^{2}\big),$$
(2.22)

where we have defined

$$\Delta_{\varepsilon}\phi(x) := 2\varepsilon^{-2} \sum_{z \in \mathbb{Z}^d} p(z) \big(\phi(x + \varepsilon z) - \phi(x)\big),$$

$$R_{\varepsilon}\phi(x) := \varepsilon^{-2} \sum_{z \in \mathbb{Z}^d} p(z) \big(\phi(x + \varepsilon z) - \phi(x)\big)^2.$$
(2.23)

It follows that

$$M_t^n(\phi) := X_t^n(\phi) - \int_0^t \mathrm{d}s \, X_s^n \left(\frac{1}{2}\Delta_{\varepsilon_n}\phi + \varepsilon_n^{-2}(b_n - d'_n)\phi\right) \tag{2.24}$$

is a square integrable martingale with predictable quadratic variation process

$$\langle M^n(\phi) \rangle_t = \int_0^t \mathrm{d}s \, X_s^n \big(\delta_n R_{\varepsilon_n}(\phi) + \delta_n \varepsilon_n^{-2} (b_n + d'_n) \phi^2 \big). \tag{2.25}$$

Since ϕ is three times continuously differentiable, it is not hard to check that

$$\Delta_{\varepsilon_n} \phi \underset{n \to \infty}{\longrightarrow} \sigma^2 \Delta \phi \quad \text{and} \quad \delta_n R_{\varepsilon_n}(\phi) \underset{n \to \infty}{\longrightarrow} 0, \tag{2.26}$$

Taking into account (2.17), this "proves" the theorem.

2.4 The log-Laplace semigroup

Although this will not be needed in the remainder of these lectures, in this section, we elaborate a bit on the question how to prove existence and uniqueness of super-Brownian motion and introduce an important tool in its study, which is its log-Laplace semigroup.

It turns out that there exists a nonlinear semigroup $(U_t)_{t\geq 0}$, acting on nonnegative measurable functions $f : \mathbb{R} \to \mathbb{R}$, such that if $(X_t)_{t\geq 0}$ is super-Brownian motion started in $X_0 = \rho$, then

$$\mathbb{E}^{\rho}[e^{-X_t(f)}] = e^{-\rho(U_t f)} \qquad (t \ge 0).$$
(2.27)

To understand this on a heuristic level, assume that $\rho = \varepsilon \sum_{i=1}^{n} \delta_{x_i}$ and let X^i be the process started in $X_0^i = \varepsilon \delta_{x_i}$. Then, by the branching property,

$$\mathbb{E}^{\rho}[e^{-X_t(f)}] = \mathbb{E}[\prod_{i=1}^n e^{-X_t^i(f)}] = \prod_{i=1}^n \mathbb{E}[e^{-X_t^i(f)}] = e^{-\rho(U_t f)}, \quad (2.28)$$

where we have defined

$$U_t f(x) := -\varepsilon^{-1} \log \mathbb{E}^{\varepsilon \delta_x} \left[e^{-X_t(f)} \right] \qquad (x \in \mathbb{R}, \ t \ge 0).$$
(2.29)

Assuming continuity of the expression in (2.27), one 'deduces' from this that the semigroup of super-Brownian motion maps 'multiplicative' functions of the form $\rho \mapsto e^{-\rho(f)}$ into functions of the same form again, only with f replaced by a different function, which we call $U_t f$.

Some more heuristic calculations convince us that $u_t(x) := U_t f(x)$ solves the semilinear Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t}u_t = \frac{1}{2}\Delta u_t + \beta u_t - \gamma u_t^2 & (t \ge 0), \\ u_0 = f. \end{cases}$$
(2.30)

On a more formal level, the arguments go the other way around. One first proves that the Cauchy problem (2.30) has a unique solution (in an appropriate sense) and then uses the martingale problem for X to show that (2.27) holds, which then implies the branching property. This argument then shows that solutions to the martingale problem (2.10)-(2.11)are unique. Existence of solutions to this martingale problem can be proved by discrete approximation, for example as in Section 2.3.

3 Convergence of sparse voter models

3.1 Voter models

Let $\{0,1\}^{\mathbb{Z}^d}$ be the space of all functions $\eta : \mathbb{Z}^d \to \{0,1\}$ and let p be a probability distribution on \mathbb{Z}^d satisfying (2.1). By definition, the *voter model* on \mathbb{Z}^d with *invasion rates* p(y-x) is the Markov process $(\eta_t)_{t\geq 0}$ with state space $\{0,1\}^{\mathbb{Z}^d}$ that evolves according to the jump rates

$$\eta \mapsto \eta + \delta_y \quad \text{with rate} \quad p(y-x) \mathbf{1}_{\{\eta(y)=0, \eta(x)=1\}} \\ \eta \mapsto \eta - \delta_y \quad \text{with rate} \quad p(y-x) \mathbf{1}_{\{\eta(y)=1, \eta(x)=0\}}.$$
(3.1)

We may interpret $\eta_t(x)$ as the *type* of the organism living at time t at site x. Another way of expressing (3.1) is that with rate p(y - x), the site y adopts the type of the site x (which has no effect if both sites have the same type).

Voter models are linear systems in the sense that if $(S_t)_{t\geq 0}$ is the semigroup of a voter model, (P_t) is the semigroup of the random walk with jump kernel p, and l_{ϕ} is a linear function as in (2.4), then

$$S_t l_\phi = l_{P_t \phi} \qquad (t \ge 0). \tag{3.2}$$

We may construct a voter model by means of a graphical representation, where for each ordered pair of sites (x, y), at times chosen according to an independent Poisson point process with intensity p(y - x), we draw an arrow from x to y meaning that at this time, the site y adopts the type of the site x. By following these arrows back in time, we can find out where the type of a given site at a given time originates from. Let $R_s^{x,t}$ be the unique 'ancestor' at time t - s of the organism living at time t at site x. Then

$$\eta_t(x) := \eta_{t-s}(R_s^{x,t}) \qquad (x \in \mathbb{Z}^d, \ 0 \le s \le t).$$
(3.3)

For fixed t > 0, the collection of \mathbb{Z}^d -valued processes $(R_s^{x,t})_{0 \le s \le t}^{x \in \mathbb{Z}^d}$ is a Markov process which evolves in such a way that at times chosen according to a Poisson point process with intensity p(x - y), all processes that are at the position x jump to the position y. In particular, for any $x \ne y$, the processes $R^{x,t}$ and $R^{y,t}$ behave as independent random walks with jump rates p(x - y) untill the time

$$\tau(x,y) = \tau_t(x,y) := \inf\{s \ge 0 : R_s^{x,t} = R_s^{y,t}\}$$
(3.4)

and from that time onwards they move together as one random walk. Thus, $(R_s^{x,t})_{0\leq s\leq t}^{x\in\mathbb{Z}^d}$ is a system of *coalescing random walks* with jump rates p(x-y). We note that in (3.4), for convenience, we have extended our processes $(R_s^{x,t})_{0\leq s\leq t}^{x\in\mathbb{Z}^d}$ to all $s\geq 0$. Since the law of $\tau_t(x,y)$ does not depend on t, we sometimes drop the subscript t.

The behavior of a voter model depends on the dimension. Note that the difference $R_s^{x,t} - R_s^{y,t}$ between two coalescing random walks with jump rates p(x - y) is a random walk with jump rates 2p(x - y) and absorption in the origin. Our assumptions on p in (2.1) imply that such a random walk is recurrent in dimensions 1 and 2 but transient in dimensions $d \ge 3$. This implies that in dimensions 1 and 2, regardless of the initial state, for any $x \neq y$:

$$\mathbb{P}[\eta_t(x) \neq \eta_t(x)] = \mathbb{P}[\eta_0(R_t^{x,t}) \neq \eta_0(R_t^{y,t})] \le \mathbb{P}[R_t^{x,t} \neq R_t^{y,t}] = \mathbb{P}[\tau(x,y) > t] \underset{t \to \infty}{\longrightarrow} 0.$$
(3.5)

This sort of behavior is described by saying that the model *clusters*.

On the other hand, in dimensions $d \ge 3$, if we start the process in an initial law such that the $(\eta_0(x))_{x\in\mathbb{Z}^d}$ are i.i.d. with $\mathbb{P}[\eta_0(x)=1]=\theta\in(0,1)$, then, for any $x\neq y$,

$$\mathbb{P}[\eta_t(x) \neq \eta_t(x)] = 2\theta(1-\theta)\mathbb{P}[\tau(x,y) > t]$$

= $2\theta(1-\theta)\mathbb{P}[\tau_0(x,y) > t] \xrightarrow[t \to \infty]{} 2\theta(1-\theta)\mathbb{P}[\tau_0(x,y) = \infty] > 0$ (3.6)

Using this, it is not hard to prove that for each $0 < \theta < 1$, the process has an invariant law in which both types coexist a.s. and the intensity of ones is θ .

3.2 Convergence in recurrent dimensions

The following interesting theorem was proved in [CDP00].

Theorem 2 (Scaling limit of sparse voter models in transient dimensions) Let ε_n be positive constants, converging to zero and set $\delta_n := \varepsilon_n^2$. Let $(\eta_t^n)_{t\geq 0}$ be voter models on \mathbb{Z}^d $(d \geq 3)$, with invasion rates p satisfying (2.1). Define $\mathcal{M}(\mathbb{R}^d)$ -valued processes $(X_t^n)_{t\geq 0}$ by

$$X_{\varepsilon_n^2 t}^n := \delta_n \sum_{x \in \mathbb{Z}^d} \eta_t^n(x) \delta_{\varepsilon_n x} \qquad (t \ge 0).$$
(3.7)

Assume that $X_0^n \Rightarrow X_0$ for some $X_0 \in \mathcal{M}(\mathbb{R}^d)$ and let $(X_t)_{t\geq 0}$ be the super-Brownian motion with this initial state and diffusion constant σ^2 , growth rate zero and activity

$$\gamma := 2 \sum_{z \in \mathbb{Z}^d} p(z) \mathbb{P}[\tau(0, z) = \infty], \tag{3.8}$$

where $\tau(x, y)$ is the first meeting time of two independent random walks with kernel p started from x and y. Then

$$\mathbb{P}[(X_t^n)_{t\geq 0}\in\cdot] \underset{n\to\infty}{\Longrightarrow} \mathbb{P}[(X_t)_{t\geq 0}\in\cdot].$$
(3.9)

Idea of the proof Since the number of particles is of order $\delta_n^{-1} = \varepsilon_n^{-2}$ while the number of lattice points per unit of space is $(\varepsilon_n)^{-d}$, the assumption $d \ge 3$ implies that we are looking at *sparse* systems, i.e., systems where only a small fraction of the sites is of type one. As a result, because of the transience of random walk, particles should not feel each other's presence too much and effectively behave as if they were independent.

Sketch of the proof Define linear functions $f_{\phi}^{\varepsilon,\delta}$ as in (2.20). Repeating our earlier calculations, but for the new generator, we find that

$$\varepsilon^{-2}Gf_{\phi}^{\varepsilon,\delta}(\eta) = \delta\varepsilon^{-2}\sum_{\substack{x,y\\ x,y}} p(y-x)\eta(x)(1-\eta(y))\phi(\varepsilon y) +\delta\varepsilon^{-2}\sum_{\substack{x,y\\ x,y}} p(y-x)(1-\eta(x))\eta(y)(-\phi(\varepsilon y)) = \delta\varepsilon^{-2}\sum_{\substack{x,y\\ x,y}} p(y-x)(\eta(x)-\eta(y))\phi(\varepsilon y) = \delta\varepsilon^{-2}\sum_{\substack{x,y\\ x,y}} p(y-x)\eta(x)(\phi(\varepsilon y)-\phi(\varepsilon x)).$$
(3.10)

where in the last step we have split our sum into two terms and changed the summation order in one term using the symmetry of p. Similarly

$$\varepsilon^{-2}\Gamma(f_{\phi}^{\varepsilon,\delta}, f_{\phi}^{\varepsilon,\delta})(\eta) = \delta^{2}\varepsilon^{-2}\sum_{x,y} p(y-x)\eta(x)(1-\eta(y))\phi(\varepsilon y)^{2} \\ + \delta^{2}\varepsilon^{-2}\sum_{x,y} p(y-x)(1-\eta(x))\eta(y)\phi(\varepsilon y)^{2} \\ = \delta^{2}\varepsilon^{-2}\sum_{x,y} p(y-x)1_{\{\eta(x)\neq\eta(y)\}}\phi(\varepsilon y)^{2} \\ = \delta^{2}\varepsilon^{-2}\sum_{x,y} p(y-x)\eta(x)(1-\eta(y))\big(\phi(\varepsilon x)^{2} + \phi(\varepsilon y)^{2}\big).$$

$$(3.11)$$

We note that the expression in (3.10) is linear in η , reflecting the 'linear' property of voter models in (3.2), but the expression in (3.11) is not linear in η . It follows from (3.10) and (3.11) that

$$M_t^n(\phi) := X_t^n(\phi) - \int_0^t \mathrm{d}s \, X_s^n\left(\frac{1}{2}\Delta_{\varepsilon_n}\phi\right) \tag{3.12}$$

is a square integrable martingale with predictable quadratic variation process

$$\langle M^n(\phi) \rangle_t = \delta_n^2 \varepsilon_n^{-2} \int_0^t \mathrm{d}s \, \sum_{x,y} p(y-x) \eta_{\varepsilon_n^{-2}s}^n(x) \left(1 - \eta_{\varepsilon_n^{-2}s}^n(y)\right) \left(\phi(\varepsilon x)^2 + \phi(\varepsilon y)^2\right). \tag{3.13}$$

The proof now proceeds by using (3.12) and (3.13) plus some additional information to show that the laws of the processes $(X_t^n)_{t\geq 0}$ are tight. Once this is proved, by going to a subsequence if necessary, we may assume that they converge weakly to the law of a process X. We are done if we can show that (for each subsequence), X is a super-Brownian motion with parameters as in the theorem. By taking the limit in (3.12), we see in the same way as in the proof of Theorem 1 that

$$M_t(\phi) := X_t(\phi) - \int_0^t \mathrm{d}s \, X_s\left(\frac{1}{2}\sigma^2 \Delta_{\varepsilon_n}\phi\right) \tag{3.14}$$

is a square integrable martingale. We claim that its quadratic variation process is given by

$$\langle M(\phi) \rangle_t = \int_0^t \mathrm{d}s \, X_s \big(\gamma \phi^2 \big).$$
 (3.15)

This will follow by taking the limit in (3.13) provided we show that the quantity in (3.13) can be approxiated by

$$\langle M^n(\phi) \rangle_t \approx \delta_n \int_0^t \mathrm{d}s \, \sum_x \eta^n_{\varepsilon_n^{-2}s}(x) \gamma \phi(\varepsilon x)^2.$$
 (3.16)

As a first step, using the facts that $\delta_n = \varepsilon_n^2$ and $\phi(\varepsilon x)^2 + \phi(\varepsilon y)^2 \approx 2\phi(\varepsilon x)^2$ if x and y are a distance of order one from each other, we may approximate the quantity in (3.13) by

$$\delta_n \int_0^t \mathrm{d}s \, \sum_{x,y} p(y-x) \eta_{\varepsilon_n^{-2}s}^n(x) \left(1 - \eta_{\varepsilon_n^{-2}s}^n(y)\right) 2\phi(\varepsilon x)^2. \tag{3.17}$$

Choose

$$1 \ll t_n \ll \varepsilon_n^{-2} \quad \text{as } n \to \infty.$$
 (3.18)

Using voter model duality, we may rewrite the quantity in (3.17) as

$$2\delta_n \int_0^t \mathrm{d}s \, \sum_{x,y} \phi(\varepsilon x)^2 p(y-x) \eta_{\varepsilon_n^{-2}s-t_n}^n (R_{t_n}^{x,\varepsilon_n^{-2}s}) \left(1 - \eta_{\varepsilon_n^{-2}s-t_n}^n (R_{t_n}^{y,\varepsilon_n^{-2}s})\right). \tag{3.19}$$

Let

$$P_t^{(2)}(x,y;x',y') = \mathbb{P}\big[R_t^{x,0} = x', \ R_t^{y,0} = y'\big].$$
(3.20)

We claim that

$$P_t^{(2)}(x, y; x', y') \approx \mathbb{P}[\tau(x, y) < \infty] P_t(x, x') \mathbf{1}_{\{x'=y'\}} + \mathbb{P}[\tau(x, y) = \infty] P_t(x, x') P_t(y, y') \quad \text{as } t \to \infty,$$
(3.21)

which comes from the fact that if two random walks started close to each other coalesce somewhere during a large time interval [0, t], then with high probability this coalescence takes place at the beginning of the time interval and any motion during this beginning phase is irrelevant for where the random walks, coalesced or not, end up at the end of the time interval. Let $(\mathcal{F}_t^n)_{t\geq 0}$ denote the filtration generated by $(\eta_t^n)_{t\geq 0}$. Using this, it can be shown that in (3.19) one is allowed to make the following approximations

$$\begin{split} \eta_{\varepsilon_{n}^{-2}s}^{n}(x) \left(1 - \eta_{\varepsilon_{n}^{-2}s}^{n}(y)\right) &\approx \mathbb{E}\left[\eta_{\varepsilon_{n}^{-2}s}^{n}(x) \left(1 - \eta_{\varepsilon_{n}^{-2}s}^{n}(y)\right) \mid \mathcal{F}_{\varepsilon_{n}^{-2}s-t_{n}}\right] \\ &= \sum_{x',y'} P_{t_{n}}^{(2)}(x,y;x',y') \eta_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(x') \left(1 - \eta_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(y')\right) \\ &\approx \mathbb{P}[\tau(x,y) = \infty] \sum_{x',y'} P_{t_{n}}(x,x') P_{t_{n}}(y,y') \eta_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(x') \left(1 - \eta_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(y')\right) \\ &= \mathbb{P}[\tau(x,y) = \infty] \left[\sum_{x',y'} P_{t_{n}}(x,x') \eta_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(x') \\ &- \sum_{x',y'} P_{t_{n}}(x,x') P_{t_{n}}(y,y') \eta_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(x') \eta_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(y')\right]. \end{split}$$
(3.22)

Here the first term on the right-hand side in the square brackets gives a contribution to the quantity in (3.19) of

$$2\delta_n \int_0^t \mathrm{d}s \sum_{x,y} \phi(\varepsilon x)^2 p(y-x) \mathbb{P}[\tau(x,y)=\infty] \sum_{x'} P_{t_n}(x,x') \eta_{\varepsilon_n^{-2}s-t_n}^n(x')$$

$$= \gamma \delta_n \int_0^t \mathrm{d}s \sum_x \phi(\varepsilon x)^2 \sum_{x'} P_{t_n}(x,x') \eta_{\varepsilon_n^{-2}s-t_n}^n(x')$$

$$\approx \gamma \delta_n \int_0^t \mathrm{d}s \sum_x \phi(\varepsilon x)^2 \eta_{\varepsilon_n^{-2}s}^n(x') = \int_0^t \mathrm{d}s X_s^n(\gamma \phi),$$

(3.23)

which converges to the quantity in (3.15).

The second term on the right-hand side in the square brackets gives a contribution to the quantity in (3.19) of

$$2\delta_n \int_0^t \mathrm{d}s \, \sum_{x,y} \phi(\varepsilon x)^2 p(y-x) \mathbb{P}[\tau(x,y) = \infty] \sum_{x',y'} P_{t_n}(x,x') P_{t_n}(y,y') \eta_{\varepsilon_n^{-2}s - t_n}^n(x') \eta_{\varepsilon_n^{-2}s - t_n}^n(y').$$
(3.24)

We need to argue that this term is small in the limit. The intuition for this is that since $t_n \gg 1$, the points x' and y' are typically so far from each other that $\eta_{\varepsilon_n^{-2}s-t_n}^n(x')$ and $\eta_{\varepsilon_n^{-2}s-t_n}^n(y')$ are free of short-distance correlations, so we are asking for the simultaneous occurrence of two sites of type one for unrelated reasons. But since our voter models are sparse, the simultaneous occurrence of two ones should be a much rarer event than the occurrence of a single one, hence the quantity in (3.24) should be aymptotically smaller than the quantity in (3.23).

To make this intuition a bit more precise, we observe that

$$\mathbb{E}\left[\sum_{x',y'} P_{t_n}(x,x') P_{t_n}(y,y') \eta_{\varepsilon_n^{-2}s-t_n}^n(x') \eta_{\varepsilon_n^{-2}s-t_n}^n(y')\right] \\
= \mathbb{E}\left[\sum_{x',y'} P_{t_n}(x,x') P_{t_n}(y,y') \eta_0^n (R_{\varepsilon_n^{-2}s-t_n}^{x',\varepsilon_n^{-2}s-t_n}) \eta_0^n (R_{\varepsilon_n^{-2}s-t_n}^{y',\varepsilon_n^{-2}s-t_n}))\right] \\
\approx \sum_{x',y'} P_{\varepsilon_n^{-2}s}(x,x') P_{\varepsilon_n^{-2}s}(y,y') \eta_0^n(x') \eta_0^n(y'),$$
(3.25)

where in the last step we have used that x' and y' are typically so far from each other that random walks started there do not coalesce. It follows that the expectation of the quantity in (3.24) is approximately given by

$$2\delta_{n} \int_{0}^{t} \mathrm{d}s \sum_{x,y} \phi(\varepsilon x)^{2} p(y-x) \mathbb{P}[\tau(x,y) = \infty] \sum_{x',y'} P_{\varepsilon_{n}^{-2}s}(x,x') P_{\varepsilon_{n}^{-2}s}(y,y') \eta_{0}^{n}(x') \eta_{0}^{n}(y')$$

$$\leq 2 \|\phi\|_{\infty}^{2} \delta_{n} \int_{0}^{t} \mathrm{d}s \sum_{x,y} p(y-x) \sum_{x',y'} P_{\varepsilon_{n}^{-2}s}(x,x') P_{\varepsilon_{n}^{-2}s}(y,y') \eta_{0}^{n}(x') \eta_{0}^{n}(y')$$

$$\approx 2 \|\phi\|_{\infty}^{2} \delta_{n} \int_{0}^{t} \mathrm{d}s \sum_{x} \sum_{x',y'} P_{\varepsilon_{n}^{-2}s}(x,x') P_{\varepsilon_{n}^{-2}s}(x,y') \eta_{0}^{n}(x') \eta_{0}^{n}(y')$$

$$= 2 \|\phi\|_{\infty}^{2} \delta_{n} \int_{0}^{t} \mathrm{d}s \sum_{x',y'} P_{2\varepsilon_{n}^{-2}s}(x',y') \eta_{0}^{n}(x') \eta_{0}^{n}(y').$$
(3.26)

Let $Q_t(x, y)$ denote the transition density of a Brownian motion with speed σ^2 . Then we may approximate the quantity in (3.26) by

$$2\|\phi\|_{\infty}^{2} \delta_{n} \int_{0}^{t} \mathrm{d}s \sum_{x',y'} \varepsilon_{n}^{d} Q_{2s}(\varepsilon_{n}x',\varepsilon_{n}y')\eta_{0}^{n}(x')\eta_{0}^{n}(y') = 2\|\phi\|_{\infty}^{2} \delta_{n}^{-1} \varepsilon_{n}^{d} \int_{0}^{t} \mathrm{d}s \int X_{0}^{n}(\mathrm{d}x) \int X_{0}^{n}(\mathrm{d}y) Q_{2s}(x,y).$$
(3.27)

Since $\delta_n = \varepsilon_n^2$ this is a quantity of order ε_n^{d-2} , which tends to zero by our assumption that $d \ge 3$.

3.3 Convergence in dimension two

In our sketch of the proof of Theorem 2, we have used several times that the dimension d is strictly larger than 2. It may therefore seem that Theorem 2 cannot be generalized to dimension two. On closer inspection, however, the only fact about coalescing random walks in dimensions $d \ge 3$ that we have essentially used is the fact that if t is large, then two walkers, started at a distance of order one of each other, with high probability either coalesce at the beginning of the time interval [0, t], or they do not coalesce at all during this whole interval. This fact is still true in dimension two. In fact, it is known that in dimension two, the coalescence time in (3.4) has the asymptotics

$$\sum_{z} p(z) \mathbb{P}[\tau(0, z) \ge t] \sim \frac{2\pi\sigma^2}{\log t} \quad \text{as } t \to \infty.$$
(3.28)

More generally, for each fixed x, y, one has $\mathbb{P}[\tau(x, y) > t] \sim c/\log t$ for some constant c > 0. Since the logarithm is a slowly varying function, this implies that $\mathbb{P}[\tau(x, y) \leq \varepsilon t | \tau(x, y) \leq t] \rightarrow 1$ as $t \rightarrow \infty$ for each $\varepsilon > 0$. In light of this, the following theorem, due to [CDP00], should perhaps not come as a total surprise.

Theorem 3 (Scaling limit of sparse voter models in dimension two) Let ε_n be positive constants, converging to zero and set $\delta_n := \varepsilon_n^2 \log(\varepsilon_n^{-2})$. Let $(\eta_t^n)_{t\geq 0}$ be voter models on \mathbb{Z}^2 with invasion rates p satisfying (2.1). Define $\mathcal{M}(\mathbb{R}^2)$ -valued processes $(X_t^n)_{t\geq 0}$ by

$$X_{\varepsilon_n^2 t}^n := \delta_n \sum_{x \in \mathbb{Z}^2} \eta_t^n(x) \delta_{\varepsilon_n x} \qquad (t \ge 0).$$
(3.29)

Assume that $X_0^n \Rightarrow X_0$ for some $X_0 \in \mathcal{M}(\mathbb{R}^2)$ and let $(X_t)_{t\geq 0}$ be the super-Brownian motion with this initial state and diffusion constant σ^2 , zero growth rate and activity $\gamma = 4\pi\sigma^2$. Then

$$\mathbb{P}[(X_t^n)_{t\geq 0}\in\cdot] \underset{n\to\infty}{\Longrightarrow} \mathbb{P}[(X_t)_{t\geq 0}\in\cdot].$$
(3.30)

Remark Since the number of particles is of order $\delta_n^{-1} = \varepsilon_n^{-2}/\log(\varepsilon_n^{-2})$ while the number of lattice points per unit of space is ε_n^{-2} , our voter models are sparse in the limit $n \to \infty$.

Sketch of the proof The proof of Theorem 2 carries over without a change up to formulas (3.12) and (3.13), and by taking the limit in (3.12) we deduce (3.14) just as in dimensions $d \geq 3$. (Note that this argument is insensitive to the precise choice of δ_n .) The approximation of (3.13) with (3.17) also remains valid in dimension two, but from this point on we need to be more careful.

In the two-dimensional case, the approximation (3.21) needs to be modified to

$$P_t^{(2)}(x, y; x', y') \approx \mathbb{P}[\tau(x, y) < t] P_t(x, x') \mathbf{1}_{\{x=x'\}} + \mathbb{P}[\tau(x, y) \ge t] P_t(x, x') P_t(y, y') \quad \text{as } t \to \infty,$$
(3.31)

which is justified by the fact that with high probability, coalescence either takes place at the beginning of the interval [0, t], or not at all during this interval (even though the walkers will a.s. eventually coalesce).

Arguing in the same way as in the proof of Theorem 2, in analogy with (3.23), one then finds that the quantity in (3.17) can be approximated by

$$2\delta_n^2 \varepsilon^{-2} \int_0^t \mathrm{d}s \, \sum_{x,y} \phi(\varepsilon x)^2 p(y-x) \mathbb{P}[\tau(x,y) \ge t_n] \sum_{x'} P_{t_n}(x,x') \eta_{\varepsilon_n^{-2}s-t_n}^n(x'). \tag{3.32}$$

(The preconstant here is different from the one in (3.23), which was derived using the fact that $\delta_n = \varepsilon_n^{-2}$ in dimensions $d \ge 3$.) Since the logarithm is a slowly varying function, by (3.28), we can choose $t_n \ll \varepsilon_n^{-2}$ but still large enough such that $\sum_y p(y-x)\mathbb{P}[\tau(x,y) \ge t_n]$ is asymptotically equivalent to

$$\sum_{y} p(y-x)\mathbb{P}[\tau(x,y) \ge \varepsilon_n^{-2}] \sim \frac{2\pi\sigma^2}{\log(\varepsilon_n^{-2})} \quad \text{as } n \to \infty.$$
(3.33)

Inserting this into (3.32), using the fact that $\delta_n = \varepsilon_n^2 \log(\varepsilon_n^{-2})$, the proof then proceeds as in the higher-dimensional case. The main work is to show that the term in (3.24) can be neglected, but this follows from the fact that we have chosen t_n large enough to get rid of local correlations and from the sparseness of our voter models, just as in dimensions $d \ge 3$.

4 Convergence of Neuhauser-Pacala models

4.1 Neuhauser-Pacala models

In [NP99], Neuhauser and Pacala introduced a model for the spatial distribution of two closely related species. Let p be a probability distribution on \mathbb{Z}^d satisfying (2.1) and assume moreover that

(v)
$$p(0) = 0.$$
 (4.1)

The Neuhauser-Pacala model with invasion rates p(y-x) and competition rates $0 \le \alpha_0, \alpha_1$ is the Markov process $(\xi_t)_{t\ge 0}$ with state space $\{0,1\}^{\mathbb{Z}^d}$ that evolves according to the jump rates

$$\xi \mapsto \xi + \delta_x \quad \text{with rate} \quad \left(f_0(x,\xi) + \alpha_0 f_0(x,\xi) \right) f_1(x,\xi)$$

$$\xi \mapsto \xi + \delta_x \quad \text{with rate} \quad \left(f_0(x,\xi) + \alpha_0 f_0(x,\xi) \right) f_1(x,\xi),$$
 (4.2)

where

$$f_i(x,\xi) := \sum_y p(y-x) \mathbf{1}_{\{\xi(y)=i\}} \qquad (i=0,1)$$
(4.3)

denotes the local frequency of type *i* in the neighborhood of the site *x*. We may interpret the rates in (4.2) as follows: an individual belonging to species *i* dies with a rate that is proportional to the number of individuals of its own species living nearby and α_i times the number of individuals of the other species living nearby. After an individual dies, it is replaced by the offspring of a randomly chosen individual living nearby. We call α_i the *competition rate* experienced by individuals of species *i* due to competition with individuals of the other species. Note that the competition rate experienced by individuals due to competition with their own species is one. The biologically relevant case is $\alpha_i \leq 1$, i.e., individuals experience less competition from the other species than from their own, due to their only partially overlapping biological niches.

If $\alpha_0 = \alpha_1 = 1$, then the model in (4.2) is a voter model. In particular, there exist no invariant laws in which the two species coexist if the dimension is one or two but there exist coexisting invariant laws in dimensions $d \geq 3$. On the other hand, assuming moreover that p(-1)+p(1) < 1 in dimension one, Neuhauser and Pacala proved that if $\alpha_0 = \alpha_1$ is sufficiently close to zero, then there exist such coexisting invariant laws in any dimension.

It is believed (but not proved) that in the symmetric case $\alpha = \alpha_0 = \alpha_1$, if there exists a coexisting invariant law for some value of α , then there exists a coexisting invariant law for all $\alpha' < \alpha$. It is moreover believed (but not proved except in the case p(-1) + p(1) = 1) that in dimension d = 1, if α is sufficiently close to one, then the model has no coexisting invariant laws. On the oter hand, due to the work of Cox, Merle and Perkins [CP05, CP07, CP08, CMP10], it is now rigorously known that in dimensions $d \ge 2$, if α is sufficiently close to one, then the model has a coexisting invariant law. If one assumes that the conjectured (but unproven) monotonicity in α is correct, then this implies that in dimensions $d \ge 2$, coexistence is possible for any $\alpha < 1$ (and in dimensions $d \ge 3$ even for $\alpha = 1$). Thus: in dimensions two and more, even the smallest difference in ecological niches is sufficient to allow species to coexist, but in dimension one species need to be sufficiently different to be able to coextist.

4.2 Convergence in transient dimensions

The following theorem is proved in [CP05] and then used in [CP07] to prove coexistence of Neuhauser-Pacala models in dimensions $d \geq 3$ for α_0, α_1 sufficiently close to each other and close to one.

Theorem 4 (Sparse Neuhauser-Pacala models in transient dimensions) Let ε_n be positive constants, converging to zero and set $\delta_n := \varepsilon_n^2$. Let $(\eta_t^n)_{t\geq 0}$ be Neuhauser-Pacala models on \mathbb{Z}^d $(d \geq 3)$, with invasion rates p satisfying (2.1) and (4.1) and competition rates

$$\alpha_i^n = 1 + \theta_i \varepsilon_n^2 \qquad (i = 0, 1), \tag{4.4}$$

where $\theta_0, \theta_1 \in \mathbb{R}$. Define $\mathcal{M}(\mathbb{R}^d)$ -valued processes $(X_t^n)_{t>0}$ by

$$X_{\varepsilon_n^2 t}^n := \delta_n \sum_{x \in \mathbb{Z}^d} \eta_t^n(x) \delta_{\varepsilon_n x} \qquad (t \ge 0).$$

$$(4.5)$$

Assume that $X_0^n \Rightarrow X_0$ for some $X_0 \in \mathcal{M}(\mathbb{R}^d)$ and let $(X_t)_{t\geq 0}$ be the super-Brownian motion with this initial state and diffusion constant σ^2 , growth rate

$$\beta = \theta_0 \sum_{z,z'} p(z)p(z')\mathbb{P}[\tau(0,z) = \tau(0,z') = \infty, \ \tau(z,z') < \infty] - \theta_1 \sum_{z,z'} p(z)p(z')\mathbb{P}[\tau(0,z) = \tau(0,z') = \infty]$$
(4.6)

and activity

$$\gamma := 2 \sum_{z \in \mathbb{Z}^d} p(z) \mathbb{P}[\tau(0, z) = \infty], \tag{4.7}$$

where $\tau(x, y)$ is defined in (3.4). Then

$$\mathbb{P}[(X_t^n)_{t\geq 0}\in\cdot]\underset{n\to\infty}{\Longrightarrow}\mathbb{P}[(X_t)_{t\geq 0}\in\cdot].$$
(4.8)

Remarks The biologically relevant case is $\theta_0, \theta_1 \leq 0$. In particular, if $\theta_0 = \theta_1 < 0$, then, since $\mathbb{P}[\tau(z, z') < \infty | \tau(0, z) = \tau(0, z') = \infty] < 1$, the growth parameter β in (4.6) is strictly positive. This reflects the fact that we are looking at models in which organisms of species 1 are sparse and the dynamics of the Neuhauser-Pacala model give an advantage to types that are locally in the minority.

Sketch of the proof The proof is similar to the proof of Theorem 2 and in fact yields the latter as a special case when $\theta_0 = \theta_1 = 0$. We may rewrite the rates in (4.2) as

$$\xi \mapsto \xi + \delta_x \quad \text{with rate} \quad f_1(x,\xi) + \varepsilon \theta_0 f_1(x,\xi)^2 \xi \mapsto \xi + \delta_x \quad \text{with rate} \quad f_0(x,\xi) + \varepsilon \theta_1 f_0(x,\xi)^2.$$

$$(4.9)$$

Writing our generator as $G_{\rm NP} = G_{\rm vot} + \varepsilon^2 G_{\theta}$, in the generator calculation (3.10), we get an additional term of the form

$$G_{\theta}f_{\phi}^{\varepsilon,\delta}(\xi) = \delta\theta_0 \sum_{x,y,z} p(y-x)p(z-x)(1-\xi(x))\xi(y)\xi(z)\phi(\varepsilon x) -\delta\theta_1 \sum_{x,y,z} p(y-x)p(z-x)\xi(x)(1-\xi(y))(1-\xi(z))\phi(\varepsilon x).$$
(4.10)

This yields two extra terms in the compensator of $X^n(\phi)$ (compare (3.12)) of the form

$$\delta_{n}\theta_{0} \int_{0}^{t} \mathrm{d}s \sum_{x,y,z} p(y-x)p(z-x)(1-\xi_{\varepsilon_{n}^{-2}s}^{n}(x))\xi_{\varepsilon_{n}^{-2}s}^{n}(y)\xi_{\varepsilon_{n}^{-2}s}^{n}(z)\phi(\varepsilon_{n}x) -\delta_{n}\theta_{1} \int_{0}^{t} \mathrm{d}s \sum_{x,y,z} p(y-x)p(z-x)\xi_{\varepsilon_{n}^{-2}s}^{n}(x)(1-\xi_{\varepsilon_{n}^{-2}s}^{n}(y))(1-\xi_{\varepsilon_{n}^{-2}s}^{n}(z))\phi(\varepsilon_{n}x).$$

$$(4.11)$$

Let $P_t^{(3)}(x, y, x; x', y', z')$ denote the transition probabilities of three coalescing random walks. We claim that (compare (3.21))

$$P_{t}^{(3)}(x, y, x; x', y', z') \approx \mathbb{P}[\tau(x, y) < \infty, \ \tau(x, z) < \infty] P_{t}(x, x') \mathbf{1}_{\{x'=y'=z'\}} \\ + \mathbb{P}[\tau(x, y) < \infty, \ \tau(x, z) = \infty] P_{t}(x, x') P_{t}(z, z') \mathbf{1}_{\{x'=y'\}} \\ + \mathbb{P}[\tau(x, z) < \infty, \ \tau(x, y) = \infty] P_{t}(x, x') P_{t}(y, y') \mathbf{1}_{\{x'=z'\}} \\ + \mathbb{P}[\tau(y, z) < \infty, \ \tau(x, y) = \infty] P_{t}(x, x') P_{t}(y, y') \mathbf{1}_{\{y'=z'\}} \\ + \mathbb{P}[\tau(x, y) = \tau(y, z) = \tau(z, x) = \infty] P_{t}(x, x') P_{t}(y, y') P_{t}(z, z').$$
(4.12)

We choose $1 \ll t_n \ll \varepsilon_n^{-2}$ and approximate (4.11) by looking a bit back in time. This yields

$$\begin{split} \delta_{n}\theta_{0} & \int_{0}^{t} \mathrm{d}s \, \sum_{x,y,z} \phi(\varepsilon_{n}x)p(y-x)p(z-x) \\ & \times \Big\{ \mathbb{P}[\tau(y,z) < \infty, \, \tau(x,y) = \infty] \sum_{x',y'} P_{t}(x,x')P_{t}(y,y')(1-\xi_{\varepsilon_{n}}^{n-2}s-t_{n}(x'))\xi_{\varepsilon_{n}}^{n-2}s-t_{n}(y') \\ & + \mathbb{P}[\tau(x,y) = \tau(y,z) = \tau(z,x) = \infty] \sum_{x',y',z'} P_{t}(x,x')P_{t}(y,y')P_{t}(z,z') \\ & \times (1-\xi_{\varepsilon_{n}}^{n-2}s-t_{n}(x'))\xi_{\varepsilon_{n}}^{n-2}s-t_{n}(y')\xi_{\varepsilon_{n}}^{n-2}s-t_{n}(z') \Big\} \\ & -\delta_{n}\theta_{1} \int_{0}^{t} \mathrm{d}s \, \sum_{x,y,z} \phi(\varepsilon_{n}x)p(y-x)p(z-x) \\ & \times \Big\{ \mathbb{P}[\tau(y,z) < \infty, \, \tau(x,y) = \infty] \sum_{x',y'} P_{t}(x,x')P_{t}(y,y')\xi_{\varepsilon_{n}}^{n-2}s-t_{n}(x')(1-\xi_{\varepsilon_{n}}^{n-2}s-t_{n}(y')) \\ & + \mathbb{P}[\tau(x,y) = \tau(y,z) = \tau(z,x) = \infty] \sum_{x',y',z'} P_{t}(x,x')P_{t}(y,y')P_{t}(z,z') \end{split}$$

$$\times \xi_{\varepsilon_n^{-2}s-t_n}^{n}(x')(1-\xi_{\varepsilon_n^{-2}s-t_n}^{n}(y'))(1-\xi_{\varepsilon_n^{-2}s-t_n}^{n}(z')) \Big\}.$$
(4.13)
ursity of ones, which says that if we look at ite x', y', z' that are sufficiently

We now use the sparsity of ones, which says that if we look at ite x', y', z' that are sufficiently far apart, then given that at least one of these positions is of type one, the other two sites are with hight probability of type zero. Skipping the more careful arguments used to justify this in previous sections, we simply approximate our formulas by

$$\delta_{n}\theta_{0}\int_{0}^{t}\mathrm{d}s\sum_{x,y,z}\phi(\varepsilon_{n}x)p(y-x)p(z-x)\mathbb{P}[\tau(y,z)<\infty,\ \tau(x,y)=\infty]\sum_{y'}P_{t}(y,y')\xi_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(y')$$
$$-\delta_{n}\theta_{1}\int_{0}^{t}\mathrm{d}s\sum_{x,y,z}\phi(\varepsilon_{n}x)p(y-x)p(z-x)\mathbb{P}[\tau(x,z)=\tau(x,y)=\infty]\sum_{x'}P_{t}(x,x')\xi_{\varepsilon_{n}^{-2}s-t_{n}}^{n}(x')$$
$$\approx\delta_{n}\beta\int_{0}^{t}\mathrm{d}s\sum_{x}\xi_{\varepsilon_{n}^{-2}s}^{n}(x)\phi(\varepsilon_{n}x)=\int_{0}^{t}\mathrm{d}s\ X_{s}^{n}(\beta\phi),$$

$$(4.14)$$

which explains the extra drift compared to the pure voter model case. The expression for the quadratic variation also gets extra terms but one can show that these are small in the limit. The arguments using the sparsety of ones become technically more demanding than in the pure voter case since duality is no longer available, even though it is still approximately true on suitably chosen intermediate time scales.

4.3 Convergence in dimension two

In [CP08], Theorem 4 is generalized to dimension 2, but the groth parameter of the limiting super-Brownian motion is not strong enough there to conclude that coexistence holds for α_0, α_1 sufficiently close to each other and close to one. This difficulty is overcome in [CMP10] where a more subtle convergence statement is proved that we formulate now.

In dimension two, one has the asymptotics (compare (3.28))

(i)
$$\sum_{z} p(z) \mathbb{P}[\tau(0,z) \ge t] \underset{t \to \infty}{\sim} \frac{2\pi\sigma^{2}}{\log t},$$

(ii)
$$\sum_{z,z'} p(z) p(z') \mathbb{P}[\tau(0,z) \ge t, \ \tau(z,z') < t] \underset{t \to \infty}{\sim} \frac{\lambda}{\log t},$$

(iii)
$$\sum_{z,z'} p(z) p(z') \mathbb{P}[\tau(0,z) \land \tau(0,z') \land \tau(z,z') \ge t] \underset{t \to \infty}{\sim} \frac{\kappa}{(\log t)^{3}},$$

(4.15)

where $0 < \lambda, \kappa < \infty$ are constants, depending on p. (In fact (i) and (iii) imply (ii) with $0 < \lambda < 2\pi\sigma^2$.)

The following result is proved in [CMP10].

Theorem 5 (Sparse Neuhauser-Pacala models in transient dimensions) Let ε_n be positive constants, converging to zero. Let $(\eta_t^n)_{t\geq 0}$ be Neuhauser-Pacala models on \mathbb{Z}^2 with invasion rates p satisfying (2.1) and (4.1) and competition rates

$$\alpha_i^n = 1 - \varepsilon_n^2 (\log(\varepsilon_n^{-2}))^3 + \theta_i \varepsilon_n^2 \log(\varepsilon_n^{-2}) \qquad (i = 0, 1),$$

$$(4.16)$$

where $\theta_0, \theta_1 \in \mathbb{R}$. Define $\mathcal{M}(\mathbb{R}^2)$ -valued processes $(X_t^n)_{t\geq 0}$ by

$$X_{\varepsilon_n^2 t}^n := \varepsilon_n^2 \log(\varepsilon_n^{-2}) \sum_{x \in \mathbb{Z}^2} \eta_t^n(x) \delta_{\varepsilon_n x} \qquad (t \ge 0).$$
(4.17)

Assume that $X_0^n \Rightarrow X_0$ for some $X_0 \in \mathcal{M}(\mathbb{R}^2)$ and let $(X_t)_{t\geq 0}$ be the super-Brownian motion with this initial state and diffusion constant σ^2 , growth rate $\beta = \kappa + \lambda(\theta_0 - \theta_1)$ and activity $\gamma = 4\pi\sigma^2$. Then

$$\mathbb{P}[(X_t^n)_{t\geq 0}\in\cdot]\underset{n\to\infty}{\Longrightarrow}\mathbb{P}[(X_t)_{t\geq 0}\in\cdot].$$
(4.18)

Remarks Contrary to Theorem 4, this theorem is only concerned with the case $\alpha_0, \alpha_1 < 1$. Note that $(\log(\varepsilon_n^{-2}))^3 \gg \log(\varepsilon_n^{-2})$, so the distance between α_0 and α_1 is much smaller than the distance between either of them and 1. This subtle way of approaching the point $(\alpha_0, \alpha_1) = (1, 1)$ is necessary in order to obtain the contribution κ to the drift, which is essential to obtain a supercritical process in the symmetric case $\alpha_0 = \alpha_1$. To see where this difficulty arises from, recall from the remarks below Theorem 4 that to obtain a positive growth rate in the symmetric case $\alpha_0 = \alpha_1$, it was essential that $\sum_{z,z'} p(z)p(z')\mathbb{P}[\tau(z,z') < \infty | \tau(0,z) = \tau(0,z') = \infty] < 1$. On the other hand (4.15) (i) and (iii) imply that in two dimensions

$$\sum_{z,z'} p(z)p(z')\mathbb{P}[\tau(z,z') < t \mid \tau(0,z) \land \tau(0,z') \ge t] \underset{t \to \infty}{\longrightarrow} 1,$$
(4.19)

whence we cannot expect a positive drift due to this effect.

Sketch of the proof One needs to adapt the proof of Theorem 4 in the same spirit as how we modified the proof of Theorem 2 to obtain Theorem 3. A difficulty is that in the calculation of the growth rate, the leading order terms cancel and one has to go to the next order. To spare the reader, we skip the details.

References

- [CDP00] J.T. Cox, R. Durrett and E.A. Perkins. Rescaled voter models converge to super-Brownian motion. Ann. Probab. 28, 185–234, 2000.
- [CMP10] J.T. Cox, M. Merle and E.A. Perkins. Coexistence in a two-dimensional Lotka-Volterra model. *Electron. J. Probab.* 15, paper No. 38, 1190-1266, 2010.
- [CP05] J.T. Cox and E.A. Perkins. Rescaled Lotka-Volterra models converge to super-Brownian motion. Ann. Probab. 33, 904–947, 2005.
- [CP07] J.T. Cox and E.A. Perkins. Survival and coexistence in stochastic spatial Lotka-Volterra models. Probab. Theory Relat. Fields 139, 89–142, 2007.
- [CP08] J.T. Cox and E.A. Perkins. Renormalization of the Two-dimensional Lotka-Volterra Model. Ann. Appl. Probab. 18, 747–812, 2008.
- [NP99] C. Neuhauser and S.W. Pacala. An explicitly spatial version of the Lotka-Volterra model with interspecific competition. Ann. Appl. Probab. 9(4): 1226–1259, 1999.