A Complete Convergence Theorem for Voter Model Perturbations

J. Theodore Cox*

Department of Mathematics
Syracuse University
e-mail: jtcox@syr.edu

Edwin A. Perkins†

Department of Mathematics
The University of British Columbia
e-mail: perkins@math.ubc.ca

Abstract: We prove a complete convergence theorem for a class of symmetric voter model perturbations with annihilating duals. A special case of interest covered by our results is the stochastic spatial Lotka-Volterra model introduced by Neuhauser and Pacala (1999). We also treat two additional models, the "affine" and "geometric" voter models.

AMS 2000 subject classifications: Primary 60K35, 82C22; Secondary 60F99.
Keywords and phrases: Complete convergence theorem, Lotka-Volterra, Interacting particle system, Voter model perturbation, Annihilating dual.

1. Introduction

In our earlier study of voter model perturbations ([5]-[9]) we found conditions for survival, extinction and coexistence for these interacting particle systems. Our goal here is to show that under additional conditions it is possible to determine all stationary distributions and their domains of attraction. We start by introducing the primary example of this work, a competition model from [27].

The state of the system at time t is represented by a spin-flip process ξ_t taking values in $\{0,1\}^{\mathbb{Z}^d}$. The dynamics will in part be determined by a fixed probability kernel $p : \mathbb{Z}^d \to [0,1]$. We assume throughout that

$$p(0) = 0,\ p(x) \text{ is symmetric, irreducible, and has covariance matrix } \sigma^2 \mathrm{I} \text{ for some } \sigma^2 \in (0,\infty). \tag{1.1}$$

For most of our results we will need to assume that $p(x)$ has exponential tails, i.e.,

$$\exists\ \kappa > 0, C < \infty \text{ such that } p(x) \le C e^{-\kappa|x|}\ \forall\ x \in \mathbb{Z}^d. \tag{1.2}$$

*Supported in part by grants from the NSA and NSF.
†Supported in part by an NSERC Discovery Grant

file: Revision-Dec-21-2012.tex date: December 21, 2012

Here $|(x_1,\dots,x_d)| = \max_i |x_i|$. We define the *local density* $f_i = f_i(x,\xi)$ of type i near $x \in \mathbb{Z}^d$ by

$$f_i(x,\xi) = \sum_{y\in\mathbb{Z}^d} p(y-x)1\{\xi(y)=i\},\ i=0,1.$$

Given $p(x)$ satisfying (1.1) and nonnegative parameters (α_0,α_1), the stochastic Lotka-Volterra model of [27], $LV(\alpha_0,\alpha_1)$, is the spin-flip process ξ_t with rate function $c_{\text{LV}}(x,\xi)$ given by

$$c_{\text{LV}}(x,\xi) = \begin{cases} f_1(x,\xi)(f_0(x,\xi)+\alpha_0 f_1(x,\xi)) & \text{if } \xi(x)=0 \\ f_0(x,\xi)(f_1(x,\xi)+\alpha_1 f_0(x,\xi)) & \text{if } \xi(x)=1. \end{cases} \tag{1.3}$$

All the spin-flip rate functions we will consider, including c_{LV}, will satisfy the hypothesis of Theorem B.3 in [25]. By that result, for such a rate function $c(x,\xi)$, there is a unique $\{0,1\}^{\mathbb{Z}^d}$-valued Feller process ξ_t with generator equal to the closure of $\Omega f(\xi) = \sum_{x\in\mathbb{Z}^d} c(x,\xi)(f(\xi^x)-f(\xi))$ on the space of functions f depending on finitely many coordinates of ξ. Here ξ^x is ξ but with the coordinate at x flipped.

One goal of [27] was to establish coexistence for $LV(\alpha_0,\alpha_1)$ for some α_i. If we let $|\xi| = \sum_{x\in\mathbb{Z}^d}\xi(x)$ and $\hat{\xi}(x) = 1-\xi(x)$ for all $x\in\mathbb{Z}^d$, then coexistence for a spin-flip process ξ_t means that there is a stationary distribution μ for ξ_t such that

$$\mu(|\xi| = |\widehat{\xi}| = \infty) = 1. \tag{1.4}$$

In [27], coexistence was proved for

$$\alpha = \alpha_0 = \alpha_1 \in [0,1) \tag{1.5}$$

close enough to 0 and $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, where

$$\mathcal{N} = \{x\in\mathbb{Z}^d : 0 < |x| \le L\}, \quad L\ge 1, \tag{1.6}$$

excluding only the case $d = L = 1$.

A special case of $LV(\alpha_0,\alpha_1)$ is the voter model. If we set $\alpha_0 = \alpha_1 = 1$ and use $f_0 + f_1 = 1$ then $c_{\text{LV}}(x,\xi)$ reduces to the rate function of the voter model,

$$c_{\text{VM}}(x,\xi) = (1-\xi(x))f_1(x,\xi) + \xi(x)f_0(x,\xi). \tag{1.7}$$

It is well known (see Chapter V of [23], Theorems V.1.8 and V.1.9 in particular) that coexistence for the voter model is dimension dependent. Let $\mathbf{0}$ (respectively, $\mathbf{1}$) be the element of $\{0,1\}^{\mathbb{Z}^d}$ which is identically 0 (respectively, 1), and let $\delta_{\mathbf{0}}$, $\delta_{\mathbf{1}}$ be the corresponding unit point masses. If $d\le 2$ then there are exactly two extremal stationary distributions, $\delta_{\mathbf{0}}$ and $\delta_{\mathbf{1}}$, and hence no coexistence. If $d\ge 3$ then there is a one-parameter family $\{P_u, u\in[0,1]\}$ of translation invariant extremal stationary distributions, where P_u has density u, i.e., $P_u(\xi(x)=1) = u$. For $u\ne 0,1$, each P_u satisfies (1.4), so there is coexistence.

Returning to the general Lotka-Volterra model, coexistence for $LV(\alpha_0, \alpha_1)$ for certain (α_0, α_1) near $(1,1)$ (including $\alpha_0 = \alpha_1 < 1$, $1 - \alpha_i$ small enough) was obtained in [8] for $d \geq 3$ and in [6] for $d = 2$. The methods used in this work require symmetry in the dynamics between 0's and 1's, i.e., condition (1.5). Under this assumption, Theorem 4 of [8] and Theorem 1.2 of [6] reduce to the following, with $LV(\alpha)$ denoting the Lotka-Volterra model when (1.5) holds.

Theorem A. *Assume $d \geq 2$ and (1.5) holds. If $d = 2$, assume also that $\sum_{x \in \mathbb{Z}^2} |x|^3 p(x) < \infty$. Then there exists $\alpha_c = \alpha_c(d) < 1$ such that coexistence holds for $LV(\alpha)$ for $\alpha \in (\alpha_c, 1)$.*

Given coexistence, one would like to know if there is more than one stationary distribution satisfying the coexistence condition (1.4), if so what are all stationary distributions, and from what initial states is there weak convergence to a given stationary distribution. To state our answers to these questions for $LV(\alpha)$ we need some additional notation. Define the hitting times

$$\tau_{\mathbf{0}} = \inf\{t \geq 0 : \xi_t = \mathbf{0}\},\ \tau_{\mathbf{1}} = \inf\{t \geq 0 : \xi_t = \mathbf{1}\},$$

and the probabilities, for $\xi \in \{0,1\}^{\mathbb{Z}^d}$,

$$\beta_0(\xi) = P_\xi(\tau_{\mathbf{0}} < \infty), \quad \beta_1(\xi) = P_\xi(\tau_{\mathbf{1}} < \infty), \quad \beta_\infty(\xi) = P_\xi(\tau_{\mathbf{0}} = \tau_{\mathbf{1}} = \infty),$$

where P_ξ is the law of our process starting at ξ. The point masses $\delta_{\mathbf{0}}, \delta_{\mathbf{1}}$ are clearly stationary distributions for $LV(\alpha)$. We write $\xi_t \Rightarrow \mu$ to mean that the law of ξ_t converges weakly to the probability measure μ. A law μ on $\{0,1\}^{\mathbb{Z}^d}$ is symmetric iff $\mu(\xi \in \cdot) = \mu(\hat{\xi} \in \cdot)$.

We note here that for any translation invariant spin-flip system ξ_t satisfying the hypothesis (B4) of Theorem B.3 in [25],

$$\beta_0(\xi) = 0 \text{ if } |\xi| = \infty, \text{ and } \beta_1(\xi) = 0 \text{ if } |\hat{\xi}| = \infty. \tag{1.8}$$

To see this for β_0, assume ξ_0 satisfies $|\xi_0| = \infty$. By assumption, there is a uniform maximum flip rate at any site in any configuration of M, so for ξ_0 and x such that $\xi_0(x) = 1$, $P(\xi_t(x) = 1) \geq e^{-Mt}$. Since $|\xi_0| = \infty$, we may choose $A_n \subset \mathbb{Z}^d$ satisfying $|A_n| = n$, $\min\{|x-y| : x, y \in A_n, x \neq y\} \to \infty$ as $n \to \infty$, and $\xi_0(x) = 1\ \forall\ x \in A_n$. Our hypotheses and translation invariance allow us to apply Theorem I.4.6 of [23] and conclude that for any fixed $t > 0$, $E\Big(\prod_{x \in A_n} \hat{\xi}_t(x)\Big) - \prod_{x \in A_n} E(\hat{\xi}_t(x)) \to 0$. It follows that for any n, there are $\{\varepsilon_n\}$ approaching 0 so that

$$\begin{aligned} P(\xi_t(x) = 0\ \forall\ x \in \xi_0) &\leq P(\xi_t(x) = 0\ \forall\ x \in A_n) \\ &\leq \varepsilon_n + \prod_{x \in A_n} P(\xi_t(x) = 0) \\ &\leq \varepsilon_n + (1 - e^{-Mt})^n \to 0, \text{ as } n \to \infty. \end{aligned}$$

Recall (see Corollary V.1.13 of [23]) that for the voter model itself and ξ_0 translation invariant with $P(\xi_0(x) = 1) = u$, we have $\xi_t \Rightarrow u\delta_{\mathbf{1}} + (1-u)\delta_{\mathbf{0}}$ if $d \le 2$, and $\xi_t \Rightarrow P_u$ if $d \ge 3$ and ξ_0 is ergodic.

Theorem 1.1. *Assume $d \ge 2$, and (1.2). There exists $\alpha_c < 1$ such that for all $\alpha \in (\alpha_c, 1)$, $LV(\alpha)$ has a translation invariant symmetric stationary distribution $\nu_{1/2}$ satisfying the coexistence property (1.4), such that for all $\xi_0 \in \{0,1\}^{\mathbb{Z}^d}$,*

$$\xi_t \Rightarrow \beta_0(\xi_0)\delta_{\mathbf{0}} + \beta_\infty(\xi_0)\nu_{1/2} + \beta_1(\xi_0)\delta_{\mathbf{1}} \text{ as } t \to \infty. \tag{1.9}$$

Theorem 1.1 is a *complete convergence theorem*, it gives complete answers to the questions raised above. The first theorem of this type for infinite particle systems was proved for the contact process in [19], where $\beta_1(\xi) = 0$ for $\xi \ne \mathbf{1}$ and $\delta_{\mathbf{1}}$ is not a stationary distribution. Our result is closely akin to the complete convergence theorem proved in [22] for the threshold voter model. (Indeed, we make use of a number of ideas from [22].) A more recent example is Theorem 4 in [28] for the $d = 1$ "rebellious voter model." For $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, $\mathcal{N}$ as in (1.6), the existence and uniqueness of $\nu_{1/2}$ in the above context follows from results in [28] and Theorem A. The relationship between Theorem 1.1 and results in [28] is discussed further in Remarks 2 and 3 below.

For $LV(\alpha)$, we note that if $0 < |\xi_0| < \infty$ then $0 < \beta_0(\xi_0) < 1$, where the upper bound is valid for α close enough to 1, and if $|\xi_0| = \infty$ then $\beta_0(\xi_0) = 0$. By the symmetry condition (1.5), this implies that the obvious symmetric statements with $(\widehat{\xi}_0, \beta_1)$ in place of (ξ_0, β_0) also hold by (1.5). To see the above, note first that $|\xi_0| < \infty$ trivially implies $\beta_0 > 0$ since one can prescribe a finite sequence of flips that leads to the trap $\mathbf{0}$. The fact that $\beta_0 < 1$ for $\alpha < 1$ close enough to 1 follows from the survival results in [8] for $d \ge 3$ (see Theorem 1 there), and in [6] for $d = 2$ (see Theorem 1.4 there). Finally, $\beta_0(\xi_0) = 0$ if $|\xi_0| = \infty$ holds by (1.8).

As our earlier comments on the ergodic theory of the voter model show, the situation is quite different for $\alpha = 1$ as (1.9) does not hold. Moreover, by constructing blocks of alternating 0's and 1's on larger and larger annuli one can construct an initial $\xi_0 \in \{0,1\}^{\mathbb{Z}^d}$ for which the law of ξ_t does not converge as $t \to \infty$. This suggests that the above theorem is rather delicate. Nonetheless we make the following conjecture:

Conjecture 1. *For $\alpha_i < 1$, close enough to 1 and with $\alpha = (\alpha_0, \alpha_1)$ in the coexistence region of Theorem 1.10 of [5], the complete convergence theorem holds with a unique non-trivial stationary distribution ν_α in place of $\nu_{1/2}$.*

If α approaches $(1,1)$ so that $\frac{1-\alpha_1}{1-\alpha_0} \to m$, then by Theorem 1.10 of [5], the limiting particle density of ν_α must approach $u^*(m)$, where u^* is as in (1.50) of [5]. Hence one obtains the one-parameter family of invariant laws for the voter model in the limit along different slopes approaching $(1,1)$.

The $d \ge 3$ case of Theorem 1.1 is a special case of a general result for certain *voter model perturbations*. We will define this class following the formulation in [5] (instead of [7]), and then give the additional required definitions needed for

our general result. A *voter model perturbation* is a family of spin-flip systems ξ_t^ε, $0 < \varepsilon \le \varepsilon_0$ for some $\varepsilon_0 > 0$, with rate functions

$$c_\varepsilon(x,\xi) = c_{\rm VM}(x,\xi) + \varepsilon^2 c_\varepsilon^*(x,\xi) \ge 0,\ x \in \mathbb{Z}^d,\ \xi \in \{0,1\}^{\mathbb{Z}^d}, \tag{1.10}$$

where $c_\varepsilon^*(x,\xi)$ is a translation invariant, signed perturbation of the form

$$c_\varepsilon^*(x,\xi) = (1-\xi(x))h_1^\varepsilon(x,\xi) + \xi(x)h_0^\varepsilon(x,\xi). \tag{1.11}$$

Here we assume (1.1) and (1.2) hold, and for some finite N_0 there is a law q_Z of $(Z^1,\dots,Z^{N_0}) \in \mathbb{Z}^{dN_0}$ and function g_i^ε on $\{0,1\}^{N_0}$, $i=0,1$, and $\varepsilon_1 \in (0,\infty]$ so that $g_i^\varepsilon \ge 0$, and for $i=0,1$, $\xi \in \{0,1\}^{\mathbb{Z}^d}$, $x \in \mathbb{Z}^d$, and $\varepsilon \in (0,\varepsilon_0]$,

$$h_i^\varepsilon(x,\xi) = -\varepsilon_1^{-2} f_i(x,\xi) + E_Z(g_i^\varepsilon(\xi(x+Z^1),\dots\xi(x+Z^{N_0}))). \tag{1.12}$$

Here E_Z is expectation with respect to q_Z. We also suppose that (decrease $\kappa > 0$ if necessary)

$$P(Z^* \ge x) \le Ce^{-\kappa x} \text{ for } x > 0, \tag{1.13}$$

where $Z^* = \max\{|Z^1|,\dots|Z^{N_0}|\}$, and there are limiting maps $g_i : \{0,1\}^{N_0} \to \mathbb{R}_+$ such that for some $c_g, r_0 > 0$,

$$\|g_i^\varepsilon - g_i\|_\infty \le c_g \varepsilon^{r_0},\ i=0,1. \tag{1.14}$$

In addition, we will always assume that for $0 < \varepsilon \le \varepsilon_0$,

$$\mathbf{0} \text{ is a trap for } \xi_t^\varepsilon,\ \text{that is, } c_\varepsilon(x,\mathbf{0}) = 0. \tag{1.15}$$

In adding (1.14) and (1.15) to the definition of *voter model perturbation* we have taken some liberty with the definition in [5] but these conditions do appear later in that work for all the results to hold.

It is easy to check that $LV(\alpha_0,\alpha_1)$ is a voter model perturbation, as is done in Section 1.3 of [5]. We will just note here that if $\alpha_i = \alpha_i^\varepsilon = 1+\varepsilon^2\theta_i$, $\theta_i \in \mathbb{R}$, and $h_i^\varepsilon(x,\xi) = \theta_{1-i} f_i(x,\xi)^2$, $i=0,1$, then $c_{\rm LV}(x,\xi)$ has the form given in (1.10) and (1.11). Additional examples of voter model perturbations are given in Section 1 of [5]. In fact, many interesting models from the life sciences and social sciences reduce to the voter model for a specific choice of parameters, and thus in many cases can be viewed as voter model perturbations.

Coexistence results for voter model perturbations are given in [5] and [8] for $d \ge 3$ (and for the two-dimensional Lotka-Volterra model in [6]). Here we will additionally require that our voter model perturbations be *cancellative processes*, which we now define following Section III.4 of [23] (see also Chapter III of [20]). Let Y be the collection of finite subsets of $\mathbb{Z}^d$ and for $x \in \mathbb{Z}^d$, $\xi \in \{0,1\}^{\mathbb{Z}^d}$ and $A \in Y$, let $H(\xi,A) = \prod_{a\in A}(2\xi(a)-1)$ (an empty product is 1). We will call a translation invariant flip rate function $c(x,\xi)$ (not necessarily a voter model perturbation) cancellative if there is a positive constant k_0 and a map $q_0 : Y \to [0,1]$ such that

$$c(x,\xi) = \frac{k_0}{2}\Big(1 - (2\xi(x)-1)\sum_{A\in Y} q_0(A-x)H(\xi,A)\Big), \tag{1.16}$$

where $A - x = \{a - x : a \in A\}$, $q_0(\emptyset) = 0$,

$$\sum_{A \in Y} q_0(A) = 1, \text{ and} \tag{1.17}$$

$$\sum_{A \in Y} |A| q_0(A) < \infty. \tag{1.18}$$

This is a subclass of the corresponding processes defined in [23]. It follows from (1.17) that $c(x, \mathbf{1}) = 0$ and so $\mathbf{1}$ is a trap for ξ. The above rate will satisfy the hypothesis of Theorem B.3 in [25] and so, as discussed above, determines a unique $\{0,1\}^{\mathbb{Z}^d}$-valued Feller process–see the discussion in Section III.4 of [23] leading to (4.8) there. (One can also check easily that the same is true of our voter model perturbations but at times we will only assume the above cancellative property.)

Given $c(x, \xi), k_0, q_0$ as above, we can define a continuous time Markov chain taking values in Y by the following. For $F, G \in Y$, $F \neq G$, define

$$Q(F, G) = k_0 \sum_{x \in F} \sum_{A \in Y} q_0(A - x) 1\{(F \setminus \{x\}) \Delta A = G\}, \tag{1.19}$$

where Δ is the symmetric difference operator. As noted in [23], Q is the Q-matrix of a nonexplosive Markov chain ζ_t taking values in Y (see also [20]). If we think of ζ_t as the set of sites occupied by a system of particles at time t, then the interpretation of (1.19) is this. If the current state of the chain is F then at rate k_0 for each $x \in F$,

1. x is removed from F, and
2. with probability $q_0(A - x)$, particles are sent from x to A, with the proviso that a particle landing on an occupied site y annihilates itself and the particle at y.

Perhaps the simplest example of a cancellative/annihilative pair (ξ_t, ζ_t) is the voter model and its dual annihilating random walk system. Here $c_{\text{VM}}(x, \xi)$ satisfies (1.16) with $k_0 = 1$, $q_0(\{y\}) = p(y)$, $q_0(A) = 0$ if $|A| > 1$ (again, see [20] and [23]). A second example, as shown in [27], is the Lotka-Volterra process, assuming (1.5) and $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, $\mathcal{N}$ satisfies (1.6) (this will be extended to our general $p(\cdot)$'s in Section 6).

The Markov chain ζ_t is the *annihilating dual* of ξ_t. The general duality equation of Theorem III.4.13 of [23] (see also Theorem III.1.5 of [20]) and [23], simplifies in the current setting to the following *annihilating duality* equation:

$$E(H(\xi_t, \zeta_0)) = E(H(\xi_0, \zeta_t)) \quad \forall\, \xi_0 \in \{0,1\}^{\mathbb{Z}^d},\ \zeta_0 \in Y. \tag{1.20}$$

In Section 2 we will recall from [20] and [23] several implications of this duality equation for the ergodic theory of ξ_t.

Let Y_e (respectively, Y_o) denote the set of $A \in Y$ with $|A|$ even (respectively, odd). We call ζ_t (or Q) *parity preserving* if

$$Q(F, G) = 0 \text{ unless } |F|, |G| \in Y_e \text{ or } |F|, |G| \in Y_o. \tag{1.21}$$

Clearly ζ_t is parity preserving iff $q_0(A) = 0$ for all $A \in Y_e$. If ζ_t is parity preserving we will call ζ_t *irreducible* if ζ_t is irreducible on Y_o and also on $Y_e \setminus \{\emptyset\}$, and $Q(A, \emptyset) > 0$ for some $A \neq \emptyset$.

One fact we need now is Corollary III.1.8 of [20]. Let $\mu_{1/2}$ be Bernoulli product measure on $\{0,1\}^{\mathbb{Z}^d}$ with density 1/2. Then under (1.15) there is a translation invariant distribution $\nu_{1/2}$ with density 1/2 such that

$$\text{if the law of } \xi_0 \text{ is } \mu_{1/2} \text{ then } \xi_t \Rightarrow \nu_{1/2} \text{ as } t \to \infty. \tag{1.22}$$

(See (2.2) below for a proof.) For a cancellative process, $\nu_{1/2}$ will always denote this measure. We note that $\nu_{1/2}$ might be $\frac{1}{2}(\delta_{\mathbf{0}} + \delta_{\mathbf{1}})$ and hence not have the coexistence property (1.4).

Theorem 1.15 of [5] gives conditions which guarantee coexistence for ξ_t^ε for small positive ε. One assumption of that result, which we will need here, requires a function f defined in terms of the voter model equilibria P_u previously introduced. For bounded functions g on $\{0,1\}^{\mathbb{Z}^d}$ write $\langle g \rangle_u = \int g(\xi) dP_u(\xi)$, and note that $\langle g(\xi) \rangle_u = \langle g(\widehat{\xi}) \rangle_{1-u}$. As in [5], define

$$f(u) = \left\langle (1 - \xi(0))c^*(0,\xi) - \xi(0)c^*(0,\xi) \right\rangle_u, \quad u \in [0,1], \tag{1.23}$$

where c^* is as in (1.11) but with g_i in place of g_i^ε. As noted in Section 1 of [5], f is a polynomial of degree at most $N_0 + 1$, and is a cubic for $LV(\alpha_0, \alpha_1)$).

We extend our earlier definitions of β_i and τ_i to general spin-flip processes ξ.

Definition (Complete Convergence). We say that *the complete convergence theorem* holds for a given cancellative process ξ_t if (1.9) holds for all initial states $\xi_0 \in \{0,1\}^{\mathbb{Z}^d}$, where $\nu_{1/2}$ is given in (1.22), and that it holds *with coexistence* if, in addition, $\nu_{1/2}$ satisfies (1.4).

Theorem 1.2. *Assume $d \geq 3$, $c_\varepsilon(x,\xi)$ is a voter model perturbation satisfying (1.2), (1.16)-(1.18), and $f'(0) > 0$. Then there exists $\varepsilon_1 > 0$ such that if $0 < \varepsilon < \varepsilon_1$ the complete convergence theorem with coexistence holds for ξ_t^ε.*

Remark 1. *As can be seen in our proof of Theorem 1.2, it is possible to drop the exponential tail condition (1.2) if the voter model perturbations are attractive, as is the case for $LV(\alpha)$ (see e.g. (8.5) with $C_{8.3} = 1$ in [8] for the latter). To do this one uses the coexistence result in Section 6 of [8] rather than that in Section 6 of [5]. In particular it follows that in Theorem 1.1 the complete convergence result holds for the Lotka-Volterra models considered there for $d \geq 3$ without the exponential tail condition (1.2). For $LV(\alpha)$ with $d = 2$ we will have to use coexistence results in [6] to derive the complete convergence results, and instead of (1.2) these results only require*

$$\sum_{x \in \mathbb{Z}^2} |x|^3 p(x) < \infty.$$

See Remark 9 in Section 6.

Theorem 1.3 of [5] states that if the "initial rescaled approximate densities of 1's" approach a continuous function v in a certain sense, then the rescaled approximate densities of ξ_t converge to the unique solution of the reaction diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\sigma^2}{2}\Delta u + f(u), \quad u_0 = v.$$

Hence the condition $f'(0) > 0$ means there is a positive drift for the local density of 1's when the density of 1's is very small and so by symmetry a negative drift when the density of 1's is close to 1. In this way we see that this condition promotes coexistence. It also excludes voter models themselves for which the complete convergence theorem fails.

We present two additional applications of Theorem 1.2.

Example 1 (Affine Voter Model). Suppose

$$\mathcal{N} \in Y \text{ is nonempty, symmetric, and does not contain the origin.} \tag{1.24}$$

The corresponding *threshold voter model* rate function, introduced in [3], is

$$c_{\text{TV}}(x,\xi) = 1\{\xi(x+y) \neq \xi(x) \text{ for some } y \in \mathcal{N}\}.$$

See Chapter II of [25] for a general treatment of threshold voter models, and [22] for a complete convergence theorem. The affine voter model with parameter $\alpha \in [0,1]$, $AV(\alpha)$, is the spin-flip system with rate function

$$c_{\text{AV}}(x,\xi) = \alpha c_{\text{VM}}(x,\xi) + (1-\alpha)c_{\text{TV}}(x,\xi), \tag{1.25}$$

where c_{VM} is as in (1.7). This model is studied in [28] with voter kernel $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, as an example of a competition model where locally rare types have a competitive advantage.

Theorem 1.3. *Assume $d \geq 3$, (1.2) holds, and $\mathcal{N}$ satisfies (1.24). There is an $\alpha_c \in (0,1)$ so that for all $\alpha \in (\alpha_c, 1)$, the complete convergence theorem with coexistence holds for $AV(\alpha)$.*

Remark 2. *It was shown in Theorem 3(a) of [28] that, excluding the case $d = 1$ and $\mathcal{N} = \{-1,1\}$, if $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, $\mathcal{N}$ as in (1.6), and coexistence holds for $LV(\alpha)$, respectively $AV(\alpha)$, for a given $\alpha < 1$, then there is a unique translation invariant stationary distribution $\nu_{1/2}$ satisfying (1.4). Hence this is true for $LV(\alpha)$ in $d \geq 2$ for $\alpha < 1$, and sufficiently close to 1, by Theorem A, and for α sufficiently small by [27]. It is also true for $AV(\alpha)$ for $\alpha = 0$ by results in [3] and [24]. The same result in [28] also shows that if, in addition, the dual satisfies a certain "non-stability" condition, then $\xi_t \Rightarrow \nu_{1/2}$ if the law of ξ_0 is translation invariant and satisfies (1.4). The complete convergence results in Theorems 1.1 and 1.3 above (which are special cases of Theorem 1.2 if $d \geq 3$) assert a stronger and unconditional conclusion for both models for α near 1.*

Example 2 (Geometric Voter Model). Let $\mathcal{N}$ satisfy (1.24). The geometric voter model with parameter $\theta \in [0,1]$, $GV(\theta)$, is the spin-flip system with rate function

$$c_{\mathrm{GV}}(x,\xi) = \frac{1-\theta^j}{1-\theta^{|\mathcal{N}|}} \text{ if } \sum_{y\in\mathcal{N}} 1\{\xi(x+y) \neq \xi(x)\} = j, \tag{1.26}$$

where the ratio is interpreted as $j/|\mathcal{N}|$ if $\theta = 1$. This *geometric* rate function was introduced in [3], where it was shown to be cancellative. As θ ranges from 0 to 1 these dynamics range from the threshold voter model to the voter model. It turns out that the geometric voter model is a voter model perturbation for θ near 1, and the following result is another consequence of Theorem 1.2.

Theorem 1.4. *Assume $d \geq 3$ and $\mathcal{N}$ satisfies (1.24). There is a $\theta_c \in (0,1)$ so that for all $\theta \in (\theta_c, 1)$, the complete convergence theorem with coexistence holds for $GV(\theta)$.*

Remark 3. (Comparison with [22] and [28]) *The emphasis in [28] was on the use of the annihilating dual to study the invariant laws and the long time behaviour of cancellative systems. A general result (Theorem 6 of [28]) gave conditions on the dual to ensure the existence of a unique translation invariant stationary law $\nu_{1/2}$ which satisfies the coexistence property (1.4) and a stronger local non-singularity property. It also gives stronger conditions on the dual under which $\xi_t \Rightarrow \nu_{1/2}$ providing the initial law is translation invariant and satisfies the above local non-singularity condition. The general nature of these interesting results make them potentially useful in a variety of settings if the hypotheses can be verified.*

In our work we focus on cancellative systems which are also voter model perturbations. The latter allowed us to construct another dual particle system in [5], and the combined use of this dual and the annihilating dual allows one to obtain a complete convergence theorem in Theorem 1.2 for small perturbations and $d \geq 3$ ($d \geq 2$ for LV in Theorem 1.1).

Theorem 1.1 of [22] gives a complete convergence theorem for the threshold voter model, the spin-flip system with rate function c_{TV} given in Example 1 above, and a complete convergence result is *established in [28] for the one-dimensional "rebellious voter model" for a sufficiently small parameter value. In both of these works, one fundamental step is to show that the annihilating dual ζ_t grows when it survives, a result we will adapt for use here (see Lemma 2.2 and the discussion following Remark 4 below). Both [22] and [28] then use special properties of the particle systems being studied to complete the proof. In Proposition 4.1 below we give general conditions under which a cancellative spin-flip system will satisfy a complete convergence theorem with coexistence. We then verify the required conditions for the voter model perturbations arising in Theorem 1.2 and Theorem 1.1.*

We conclude this section with a "flow chart" of the proof of the main results, including an outline of the paper. First, the rather natural condition we impose that $\mathbf{0}$ is a trap for our cancellative systems ξ_t will imply that ξ_t is in fact symmetric with respect to interchange of 0's and 1's (see Lemma 2.1 in Section 2).

This helps explain the asymmetric *looking* condition $f'(0) > 0$ in Theorem 1.2 and the restriction of our results to LV with $\alpha_1 = \alpha_2$. Section 2 also reviews the ergodic theory of cancellative and annihilating systems.

As noted above, the core of our proof, Proposition 4.1, establishes a complete convergence theorem for cancellative particle systems (where $\mathbf{0}$ is a trap), assuming three conditions: (i) growth of the dual system when it survives, i.e. (2.7), (ii) a condition (4.1) ensuring a large number of 0-1 pairs at locations separated by a fixed vector x_0 for large t with high probability (ruling out clustering which clearly is an obstruction to any complete convergence theorem), and (iii) a condition (4.2) which says if the initial condition ξ_0 contains a large number of 0-1 pairs with 1's in a set A then at time 1 the probability of an odd number of 1's in A will be close to 1/2. With these inputs, the proof of Propostion 4.1 in Section 4 is a reasonably straightforward duality argument. This result requires no voter perturbation assumptions and may therefore have wider applicability.

We then verify the three conditions for voter model perturbations. The dual growth condition (2.7) is established in Lemma 2.2 and Remark 4 in Section 2, assuming the dual is irreducible and the cancellative system itself satisfies $\limsup_{t\to\infty} P(\xi_t(0) = 1) > 0$ when $\xi_0 = \delta_0$. The latter condition will be an easy by-product of our percolation arguments in Section 5. The irreducibility of the annihilating dual is proved for cancellative systems which are voter model perturbations in Section 3 (see Corollary 3.3). Condition (4.2) is verified for voter model perturbations in Lemma 4.2 of Section 4, following ideas in [1]. In Section 5 (see Lemma 5.3) the condition (4.1) is derived for the voter model perturbations in Theorem 1.2 using a comparison with oriented percolation which in turn relies on input from [5] (see Lemma 5.2) and our condition $f'(0) > 0$. Another key in this argument is the use of certain irreducibility properties of voter perturbations to help set up the appropriate block events. More specifically, with positive probability it allows us to transform a 0-1 pair at a couple of input sites at one time into a mixed configuration which has a "positive density" of both 0's and 1's at a later time (see Lemma 3.4). The percolation comparison will provide a large number of the inputs and the mixed configuration will be chosen to ensure a 0-1 pair at sites with the prescribed separation by x_0. In Section 5 we finally prove Theorem 1.2. Theorem 1.1 is proved in Section 6, and the proofs of Theorems 1.3 and 1.4 are given in Section 7. All of these latter results are proved as Corollaries to Theorem 1.2, except for the two-dimensional case of Theorem 1.1, where the input for the percolation argument is derived from [8] instead of [5].

2. Cancellative and annihilating processes: growth of the annihilating dual.

Our main objective in this Section (Lemma 2.2 below) is to show the dual growth condition: under appropriate hypotheses, the annihilating dual process ζ_t will either die out or grow without bound as $t \to \infty$.

We begin by pointing out the consequences of the assumption that $\mathbf{0}$ is a trap for ξ_t. We assume here that $c(x,\xi)$ is a translation invariant cancellative flip rate function satisfying (1.16)-(1.18), ξ_t is the corresponding cancellative process and ζ_t the corresponding annihilating process (the Markov chain on Y with Q-matrix defined in (1.19)). In part (iv) below we identify ξ_t with the set of sites of type 1. Recall that $H(\xi,A)=\prod_{a\in A}(2\xi(x)-1)$.

Lemma 2.1. *If ξ_t and ζ_t are as above, then the following are equivalent.*
(i) $\mathbf{0}$ is a trap for ξ_t.
(ii) $q_0(A)=0$ for all $A\in Y_e$, i.e., ζ_t is parity-preserving.
(iii) ξ_t is symmetric, i.e., $c(x,\xi)=c(x,\widehat{\xi})$.
(iv) The simplified duality equation holds:

$$P(|\xi_t\cap\zeta_0| \text{ is odd})=P(|\xi_0\cap\zeta_t| \text{ is odd}) \quad \forall\ \xi_0\in\{0,1\}^{\mathbb{Z}^d},\ \zeta_0\in Y. \tag{2.1}$$

Proof. Note that $H(\mathbf{0},A)=(-1)^{|A|}$, which by (1.16) implies

$$c(0,\mathbf{0})=\frac{k_0}{2}\Big(1+\sum_{A\in Y}q_0(A)(-1)^{|A|}\Big).$$

Thus $\mathbf{0}$ is a trap for ξ_t iff $\sum_{A\in Y}q_0(A)(-1)^{|A|}=-1$. Using (1.17), we see that

$$\sum_{A\in Y}q_0(A)(-1)^{|A|}=\sum_{A\in Y_e}q_0(A)-\sum_{A\in Y_o}q_0(A)\geq\sum_{A\in Y_e}q_0(A)-1,$$

so (i) and (ii) are equivalent.

Using $H(\widehat{\xi},A)=(-1)^{|A|}H(\xi,A)$ and (1.16), (ii) implies (iii) because

$$\begin{aligned}c(x,\widehat{\xi})&=\frac{k_0}{2}\Big(1-(1-2\xi(x))\sum_{A\in Y_o}q_0(A-x)(-1)^{|A|}H(\xi,A)\Big)\\&=\frac{k_0}{2}\Big(1-(2\xi(x)-1)\sum_{A\in Y}q_0(A-x)H(\xi,A)\Big)\\&=c(x,\xi).\end{aligned}$$

Conversely, if $c(0,\xi)=c(0,\widehat{\xi})$ for all ξ, the previous calculation shows that

$$\sum_{A\in Y}q_0(A)H(\xi,A)=\sum_{A\in Y}q_0(A)(-1)^{|A|+1}H(\xi,A).$$

Plug in $\xi=\mathbf{1}$ to get

$$\sum_{A\in Y}q_0(A)=\sum_{A\in Y_o}q_0(A)-\sum_{A\in Y_e}q_0(A),$$

which implies $q_0(A)=0$ if $|A|$ is even. We now have conditions (i)-(iii) are equivalent.

The duality equation (1.20) is easily seen to be equivalent to

$$P(|\zeta_0| - |\xi_t \cap \zeta_0| \text{ is odd}) = P(|\zeta_t| - |\xi_0 \cap \zeta_t| \text{ is odd}) \quad \forall\, \xi_0 \in \{0,1\}^{\mathbb{Z}^d},\ \zeta_0 \in Y.$$

If ζ_t is parity preserving then this is equivalent to (iv). Conversely, if (iv) holds, and we apply it with $\xi_0 = \mathbf{0}$ and $\zeta_0 = \{x\}$, we get $P(\xi_t(x) = 1) = 0$ for all $t > 0$. Since this holds for all $x \in \mathbb{Z}^d$, $\mathbf{0}$ must be a trap for ξ_t. □

We give a brief review (cf. [20], [23]) of the application of annihilating duality to the ergodic theory of ξ_t. Recall that $\mu_{1/2}$ is Bernoulli product measure with density 1/2 on $\{0,1\}^{\mathbb{Z}^d}$. Let ζ_t^A denote the Markov chain ζ_t with initial state A, and let ξ_0 have law $\mu_{1/2}$. It is easy to see by integrating (2.1) with respect to the law of ξ_0 that

$$\begin{aligned} P(|\xi_t \cap A| \text{ is odd}) &= E(P(|\zeta_t^A \cap \xi_0| \text{ is odd} \mid \zeta_t^A) 1(\zeta_t^A \neq \emptyset)) \\ &= \tfrac{1}{2} P(\zeta_t^A \neq \emptyset) \text{ for all } A \in Y. \end{aligned} \tag{2.2}$$

The right-hand side above is monotone in t ($\emptyset$ is a trap for ζ_t) and so the left-hand side above converges as $t \to \infty$. By inclusion-exclusion arguments the class of functions

$$\{\xi \to 1(|\xi \cap A| \text{ is odd}) : A \in Y\} \text{ is a determining class,} \tag{2.3}$$

and hence also a convergence determining class since the state space is compact. Therefore the above convergence not only implies (1.22), it characterizes $\nu_{1/2}$ via: for all $A \in Y$,

$$\nu_{1/2}(\xi : |\xi \cap A| \text{ is odd}) = \tfrac{1}{2} P(\zeta_t^A \neq \emptyset\ \forall\, t \geq 0). \tag{2.4}$$

The measure $\nu_{1/2}$ is necessarily a translation invariant stationary distribution for ξ_t with density 1/2, and a consequence of (2.4) is that $\nu_{1/2} \neq \frac{1}{2}(\delta_0 + \delta_1)$ iff for some $x \neq y \in \mathbb{Z}^d$,

$$P(\zeta_t^{\{x,y\}} \neq \emptyset\ \forall\, t \geq 0) > 0. \tag{2.5}$$

Thus, a sufficient condition for coexistence for ξ_t is (2.5). Indeed, if (2.5) holds then $\nu_{1/2}(\xi \in \cdot | \xi \notin \{\mathbf{0}, \mathbf{1}\})$ is a translation invariant stationary distribution for ξ_t which must satisfy (1.4). (There are countably many configurations ξ with $|\xi| < \infty$, none of which can have positive probability because there are countably many distinct translates of each one.)

Establishing (2.5) directly is a difficult problem for most annihilating systems. (Not so for the annihilating dual of the voter model, since (2.5) follows trivially from transience if $d \geq 3$ but fails if $d \leq 2$.) To use annihilating duality to go beyond (1.22) requires more information about the behavior of ζ_t. In particular, one needs that either $|\zeta_t| \to 0$ or $|\zeta_t| \to \infty$ as $t \to \infty$ (see [1], for instance). The following general result gives a condition for this which we can check for certain

voter model perturbations. It is a key ingredient in the proofs of Theorem 1.1 and Thereom 1.2.

We now assume that $\mathbf{0}$ is a trap for ξ_t, and so all the properties listed in Lemma 2.1 will hold.

Lemma 2.2 (Handjani [22], Sturm and Swart [28])**.** *Let ζ_t be a translation invariant annihilating process with Q-matrix given in* (1.19) *satisfying* (1.17) *and* (1.18)*. If ζ_t is irreducible, parity-preserving, and satisfies*

$$\limsup_{t\to\infty} P(0 \in \zeta_t^{\{0\}}) > 0, \tag{2.6}$$

then

$$\lim_{t\to\infty} P(0 < |\zeta_t^B| \leq K) = 0 \text{ for all non-empty } B \in Y \text{ and } K \geq 1. \tag{2.7}$$

Remark 4. *If ζ_t has associated cancellative process ξ_t which has $\mathbf{0}$ as a trap then the parity-preserving hypothesis in the above result follows by Lemma 2.1. If we let $\xi_t^{\{0\}}$ denote this process with initial state $\xi_0^{\{0\}} = \{0\}$, then by the duality equation* (2.1)*,* (2.6) *is equivalent to*

$$\limsup_{t\to\infty} P(\xi_t^{\{0\}}(0) = 1) > 0. \tag{2.8}$$

The limit (2.7) was proved in [22] (see Proposition 2.6 there) for the annihilating dual of the threshold voter model. The arguments in that work are in fact quite general, and with some work can be extended to establish Lemma 2.2 as stated above. Rather than provide the necessary details, we appeal instead to Theorem 12 of [28], which is proved using a related but somewhat different approach. To apply this result, and hence establish Lemma 2.2, we must do two things. The first is to show that (3.54) in [28] (see (2.17) below) holds, the second is to show that our condition (2.6) implies the non-stability condition in Theorem 12 of [28]. The latter is non-positive recurrence of ζ "modulo translations" (see the conclusion of Lemma 2.4 below).

In preparation for these tasks we give a "graphical construction" (as in [19] or [3]) of ζ_t. For $x \in \mathbb{Z}^d$, let $\{(S_n^x, A_n^x,) : n \in \mathbb{N}\}$ be the points of independent Poisson point processes $\{\Gamma^x(ds, dA) : x \in \mathbb{Z}^d\}$ on $\mathbb{R}_+ \times Y$ with rates $k_0 ds q_0(dA)$. For $R \subset \mathbb{R}^d$ and $0 \leq t_1 \leq t_2$ we let

$$\mathcal{F}(R \times [t_1, t_2]) = \sigma(\Gamma^x\big|_{\mathbb{Z}^d \times [t_1,t_2]} : x \in R).$$

Then for $S_i = R_i \times I_i$ as above $(i = 1, 2)$, $\mathcal{F}(S_1)$ and $\mathcal{F}(S_2)$ are independent if $S_1 \cap S_2 = \emptyset$. At time S_n^x draw arrows from x to $x + y$ for each $y \in A_n^x \setminus \{0\}$. If $0 \notin A_n^x$ put a δ at x (at time S_n^x). For $x, y \in \mathbb{Z}^d$ and $s < t$ we say that $(x, s) \to (y, t)$ if there is a path from (x, s) to (y, t) that goes across arrows, or up but not through δ's. That is, $(x, s) \to (y, t)$ if there are sequences $x_0 = x, x_1, \ldots, x_n = y$ and $s_0 = s < s_1 < \cdots < s_n < s_{n+1} = t$ such that

(i) for $1 \leq m \leq n$ there is an arrow from x_{m-1} to x_m at time s_m,

(ii) for $1 \le m \le n+1$ there are no δ's in $\{x_{m-1}\} \times (s_{m-1}, s_m)$,

and no δ at (y,t). For $0 \le s < t$, $x, y \in \mathbb{Z}^d$ and $B \in Y$ define

$$\begin{aligned} N_t^{(x,s)}(y) &= \text{ the number of paths up from } (x,s) \text{ to } (y,t), \\ \zeta_t^{B,s} &= \{y : \sum_{x \in B} N_t^{(x,s)}(y) \text{ is odd}\}, \\ \bar{\zeta}_t^{B,s} &= \{y : \sum_{x \in B} N_t^{(x,s)}(y) \ge 1\}, \end{aligned}$$

and write ζ_t^B for $\zeta_t^{B,0}$ and $\bar{\zeta}_t^B$ for $\zeta_t^{B,0}$.

The process ζ_t is the annihilating Markov chain on Y with Q-matrix as in (1.19). The process $\bar{\zeta}_t$ is additive, meaning $\bar{\zeta}_t^{B,s} = \bigcup_{x \in B} \bar{\zeta}_t^{(x,s)}$. Both ζ_t and $\bar{\zeta}_t$ are nonexplosive Markov chains on Y. Also, it is clear that for every $B \in Y$,

$$\zeta_t^{B,s} \subset \bar{\zeta}_t^{B,s} \ \forall\ 0 \le s \le t < \infty, \tag{2.9}$$

and also that for any fixed $t > 0$,

$$\lim_{K \to \infty} P(\bar{\zeta}_u^{\{0\}} \subset [-K,K]^d \ \forall\ 0 \le u \le t) = 1. \tag{2.10}$$

Furthermore if $A, B \in Y$ satisfy $\min_{a \in A, b \in B} |a-b| > 2K$ and $t > s \ge 0$, then

$\zeta_t^{A \cup B,s} = \zeta_t^{A,s} \cup \zeta_t^{B,s}$ on the event

$$\left\{\bar{\zeta}_u^{A,s} \subset A + [-K,K]^d, \bar{\zeta}_u^{B,s} \subset B + [-K,K]^d \ \ \forall s \le u \le t\right\} \tag{2.11}$$

(where $A + B = \{x + y : x \in A, y \in B\}$).

The following result is key to verifying condition (3.54) of [28].

Lemma 2.3. *Let $A \in Y$, $r \in \mathbb{N}$ and $B_m = \{y_1^m, \ldots, y_r^m\} \in Y$ be such that $\lim_{m \to \infty} \min_i |y_i^m| = \infty$. If ζ^A and ζ^{B_m} are independent copies of ζ with the given initial conditions, then for each $t \ge 0$ and $n \in \mathbb{N}$,*

$$\lim_{m \to \infty} P(|\zeta_t^{A \cup B_m}| = n) - P(|\zeta_t^A| + |\zeta_t^{B_m}| = n) = 0.$$

Proof. Assume (ζ_t^B) are constructed as above for $B \in Y$ and $t \ge 0$. For $K \in \mathbb{N}$ define $\tilde{\zeta}_t^{B,(K)}$ as ζ_t^B but now only count paths which are contained in $B + [-K,K]^d$. This implies that

$$\tilde{\zeta}_t^{B,(K)} \text{ is } \mathcal{F}((B + [-K,K]^d) \times [0,t]) - \text{measurable}. \tag{2.12}$$

Fix $\varepsilon > 0$. By (2.10) and the additivity of $\bar{\zeta}_t$ we may choose $K(\varepsilon) \in \mathbb{N}$ so that if $K \ge K(\varepsilon)$, then

$$\begin{aligned} &P(\bar{\zeta}_u^A \subset A + [-K,K]^d \text{ and } \bar{\zeta}_u^{B_m} \subset B_m + [-K,K]^d \text{ for all } u \in [0,t]) \\ &\quad > 1 - \varepsilon \ \text{ for all } m \in \mathbb{N}. \end{aligned} \tag{2.13}$$

Write $\tilde{\zeta}_t^B$ for $\tilde{\zeta}_t^{B,(K(\varepsilon))}$. Choose $m(\varepsilon) \in \mathbb{N}$ so that $\min_{a\in A, b\in B_m} |a-b| > 2K(\varepsilon)$ for $m \geq m(\varepsilon)$. It follows from (2.11) and (2.9) that on the set in (2.11) with $K = K(\varepsilon)$, for $m \geq m(\varepsilon)$,

$$|\zeta_t^{A\cup B_m}| = |\zeta_t^A| + |\zeta_t^{B_m}|, \tag{2.14}$$

and

$$\zeta_t^A = \tilde{\zeta}_t^A \text{ and } \zeta_t^{B_m} = \tilde{\zeta}_t^{B_m}. \tag{2.15}$$

(The latter is an easy check using (2.11).) We conclude from the last two results that

$$P(|\zeta_t^{A\cup B_m}| \neq |\tilde{\zeta}_t^A| + |\tilde{\zeta}_t^{B_m}|) < \varepsilon \text{ for } m \geq m(\varepsilon). \tag{2.16}$$

By (2.12) and the choice of $m(\varepsilon)$ we see that $\tilde{\zeta}_t^A$ and $\tilde{\zeta}_t^{B_m}$ are independent for $m \geq m(\varepsilon)$. Using this independence and then (2.16) we conclude that

$$\begin{aligned}
&\Big|P(|\zeta_t^{A\cup B_m}| = n) - \Big(\sum_{k=0}^n P(|\zeta_t^A| = k)P(|\zeta_t^{B_m}| = n-k)\Big)\Big| \\
&\leq \varepsilon + \Big|P(|\tilde{\zeta}_t^A| + |\tilde{\zeta}_t^{B_m}| = n) - \Big(\sum_{k=0}^n P(|\zeta_t^A| = k)P(|\zeta_t^{B_m}| = n-k)\Big)\Big| \\
&\leq \varepsilon + \Big|\sum_{k=0}^n [P(|\tilde{\zeta}_t^A| = k)P(|\tilde{\zeta}_t^{B_m}| = n-k) - P(|\zeta_t^A| = k)P(|\zeta_t^{B_m}| = n-k)]\Big| \\
&\leq 3\varepsilon.
\end{aligned}$$

In the last line we have used (2.15). The result follows. □

Say that $A, B \in Y_o$ are equivalent if they are translates of each other, let $\tilde{Y}_o$ denote the set of equivalence classes, and (abusing notation slightly) let $\tilde{A}$ denote the equivalence class containing $A \in Y$. Since the dynamics of ζ are translation invariant, for parity-preserving ζ we may define $\tilde{\zeta}_t$ as the $\tilde{Y}_o$-valued Markov process obtained by taking the equivalence class of ζ_t. The non-stability requirement of Theorem 12 of [28] is that $\tilde{\zeta}_t$ not be positive recurrent on $\tilde{Y}_o$.

Lemma 2.4. *If ζ is parity-preserving, irreducible and satisfies (2.6), then $\tilde{\zeta}_t$ is not positive recurrent.*

Proof. We use the same arguments as in the proof of Lemma 2.4 of [22]. First, ζ_t cannot be positive recurrent on Y_o. To check this, we first note that translation invariance implies

$$P(\zeta_t^{\{x\}} = \{x\}) = P(\zeta_t^{\{0\}} = \{0\}) \text{ for all } t \geq 0, x \in \mathbb{Z}^d.$$

If ζ_t is positive recurrent on Y_o then the limit $\mu(A) = \lim_{t\to\infty} P(\zeta_t^B = A)$ exists and is positive for all $A, B \in Y_o$. Letting $t \to \infty$ above, this implies

$\mu(\{0\}) = \mu(\{x\})$ for all x which is impossible, so ζ_t is not positive recurrent on Y_o. A consequence of this is that for any fixed $k > 0$,

$$\lim_{t\to\infty} P(\zeta_t^{\{0\}} \subset [-k,k]^d) = 0.$$

Next, suppose $\tilde{\zeta}_t$ is positive recurrent on $\tilde{Y}_o$, with some stationary distribution $\tilde{\mu}$, which must satisfy:

$$\tilde{\mu}(A \in \tilde{Y}_o : \mathrm{diam}(A) \le k) \to 1 \text{ as } k \to \infty,$$

where $\mathrm{diam}(A) = \max\{|x-y| : x, y \in A\}$ is well-defined for $A \in \tilde{Y}_o$. For any k, t, since $\mathrm{diam}(\tilde{\zeta}_t^{\{0\}}) = \mathrm{diam}(\zeta_t^{\{0\}})$, we have

$$P(0 \in \zeta_t^{\{0\}}) \le P(\zeta_t^{\{0\}} \subset [-k,k]^d) + P(\mathrm{diam}(\tilde{\zeta}_t^{\{0\}}) > k).$$

Letting $t \to \infty$ gives

$$\limsup_{t\to\infty} P(0 \in \zeta_t^{\{0\}}) \le \tilde{\mu}(A \in \tilde{Y}_o : \mathrm{diam}(A) > k).$$

The right-hand side above tends to 0 as $k \to \infty$, so we have a contradiction to the assumption (2.6). □

Proof of Lemma 2.2. Thanks to the above Lemma we have verified all the hypotheses of Theorem 12 of [28] except for their (3.54) which we now state in our notation: for each $n \in \mathbb{Z}_+$, $L \ge 1$, and $t > 0$,

$$\inf\{P(|\zeta_t^A| = n) : |A| = n+2 \text{ and } 0 < |i-j| \le L \text{ for some } i,j \in A\} > 0. \tag{2.17}$$

Assume (2.17) fails. Then for some n, L and t as above, by translation invariance and compactness of Y (with the subspace topology it inherits from $\{0,1\}^{\mathbb{Z}^d}$), there are $\{A_m\} \subset Y$ so that for some integer $2 \le s \le n+2$ and $x_2 \in [-L,L]^d$, $A_m = \{0, x_2, \ldots, x_s\} \cup \{x_{s+1}^m, \ldots, x_{n+2}^m\} \equiv A \cup B_m$, where $\lim_{m\to\infty} |x_i^m| = \infty$ for each $i \in \{s+1, \ldots, n+2\}$ and

$$\lim_{m\to\infty} P(|\xi_t^{A_m}| = n) = 0. \tag{2.18}$$

By the irreducibility of ζ, $P(|\zeta_t^A| = s-2) = p > 0$. If ζ_t^A and $\zeta_t^{B_m}$ are as in Lemma 2.3, then by that result,

$$\begin{aligned}\lim_{m\to\infty} P(|\zeta_t^{A_m}| = n) &= \lim_{m\to\infty} P(|\zeta_t^A| + |\zeta_t^{B_m}| = n)\\ &\ge P(|\zeta_t^A| = s-2) \liminf_{m\to\infty} P(|\zeta_t^{B_m}| = |B_m|)\\ &\ge p \exp\{-k_0(n+2-s)T\} > 0.\end{aligned}$$

In the last line we use the fact that by its graphical construction, ζ^{B_m} will remain constant up to time t if none of the $|B_m|$ independent rate k_0 Poisson processes attached to each of the sites in B_m fire by time t. This contradicts (2.18) and so (2.17) must hold. We now may apply Theorem 12 of [28] to obtain the required conclusion. □

3. Irreducibility

In addition to the explicit irreducibility requirement for ζ_t in Lemma 2.2, some arguments in Section 5 will require irreducibility type conditions for the voter model perturbations ξ_t^ε. We collect and prove the necessary results for both processes in this Section.

Assuming $\sum_{y\in\mathbb{Z}^d} q_0(\{y\}) > 0$, define the step distribution of a random walk associated with q_0 by

$$q(x) = q_0(\{x\}) \Big/ \sum_{y\in\mathbb{Z}^d} q_0(\{y\}).$$

Lemma 3.1. *Let ζ_t be a parity-preserving annihilating process with Q-matrix given in (1.19). Assume $q_0(A_0) > 0$ for some $A_0 \in Y$ with $|A_0| \geq 3$, and for some symmetric, irreducible random walk kernel r on $\mathbb{Z}^d$, $q(x) > 0$ whenever $r(x) > 0$. Then ζ_t is irreducible.*

Proof. The proof is elementary but awkward, so we will only sketch the argument. Note that if $x \in A$ and $y \notin A$ then

$$Q(A, (A \setminus \{x\}) \cup \{y\}) \geq q_0(\{y - x\}) = cq(y - x).$$

So by using only the $q_0(\{x\})$ "clocks"with $r(x) > 0$, ζ_t can with positive probability execute exactly any finite sequence of transitions that the annihilating random walk system with step distribution r can. We will refer to "r-random walks" below in describing such transitions.

We first check that the assumptions on q_0 imply that ζ_t can reach any set B with $|B| = |\zeta_0|$ with positive probability. To see this, we first construct a set B' by starting r-random walks at each site of B and then moving them apart, one at a time, avoiding collisions, to widely separated locations, resulting in B'. Note that by reversing this entire sequence of steps, it is possible to move r-random walks starting at the sites of B' to B without collisions. This uses the symmetry of r. Now, to move r-walks from ζ_0 to B we first move walks from ζ_0 to some ζ_0', avoiding collisions, where the sites of ζ_0' are widely separated. Pair off points from ζ_0' and B' and move r-walks one at a time from ζ_0' to B' without collisions. This is possible if ζ_0' and B' are sufficiently spread out since r is irreducible. Finally, move the walks from B' to B without collisions as discussed above.

It should be clear that if $\zeta_0 \neq \emptyset$ then ζ_t can reach a set B such that $|B| = |\zeta_0| - 2$, since this is the case for annihilating random walks. Finally, if $\zeta_0 \neq \emptyset$ then ζ_t can reach a set B with $|B| \geq |\zeta_0| + 2$. Choose x_1 far from ζ_0 so that ζ_0 and $x_1 + A_0$ are disjoint, and such that for some $x_0 \in \zeta_0$, an r-walk starting at x_0 can reach x_1 by a sequence of steps avoiding ζ_0. Now using the "A_0 clock" at x_1 we get a transition from $(\zeta_0 \setminus \{x_0\}) \cup \{x_1\}$ to $\zeta_0' = (\zeta_0 \setminus \{x_0\})\Delta(x_1 + A_0)$, and $|\zeta_0'| \geq |\zeta_0| + 2$. □

The next result will allow us to apply the above Lemma to voter model perturbations. Recall $p(x)$ satisfies (1.1) and $c_{\text{VM}}(x, \xi)$ is the corresponding voter model flip rate function.

Lemma 3.2. *There is an $\varepsilon_2 = \varepsilon_2(p(\cdot)) > 0$ and $R_1 = R_1(p(\cdot))$ such that*

(i) $p(\cdot \,\big|\, |x| < R_1)$ is irreducible, and

(ii) if $c(x,\xi) = c_{VM}(x,\xi) + \tilde{c}(x,\xi)$ is a translation invariant, cancellative flip rate function with $\boldsymbol{0}$ as a trap such that

$$\|\tilde{c}\|_\infty < \varepsilon_2, \quad \sum_{x\neq 0} |\tilde{c}(0,\delta_x)| < \varepsilon_2, \tag{3.1}$$

then the dual kernel q_0 satisfies

$$q_0(\{x\}) > (k_0 3)^{-1} p(x) \text{ for all } 0 < |x| < R_1. \tag{3.2}$$

Proof. Since p is irreducible, we may choose R_1 so that $p(\cdot \,|\, |x| < R_1)$ is also irreducible. Assume (3.1) holds for an appropriate ε_2 which will be chosen below. We will write $\hat{\xi}(B)$ for $\sum_{x\in B} \hat{\xi}(x)$. Also, *in this proof only,* we will let A denote a *random set* with probability mass function q_0, and write $E_0(g(A)) = \int g\, dP_0$ for $\sum_{B\in Y} g(B) q_0(B)$. With this notation, by our hypotheses we have

$$c_{\mathrm{VM}}(0,\xi) + \tilde{c}(0,\xi) = \frac{k_0}{2}\Big[1 + (-1)^{\xi(0)} E_0\Big((-1)^{\hat{\xi}(A)}\Big)\Big]. \tag{3.3}$$

Recall by Lemma 2.1 that $P_0(|A| \text{ is odd}) = 1$. Therefore if we set $\xi = \delta_x$ for $x \neq 0$ in (3.3), we get

$$p(x) + \tilde{c}(0,\delta_x) = \frac{k_0}{2}\Big[1 + E_0\Big((-1)^{|A\setminus\{x\}|}\Big)\Big] = k_0 P_0(x \in A),$$

and so

$$P_0(x \in A) = (p(x) + \tilde{c}(0,\delta_x)) k_0^{-1}. \tag{3.4}$$

If we take $\xi = \delta_{\{x_0,x_1\}}$ in (3.3), where x_0, x_1 are two distinct non-zero points then we get

$$\begin{aligned} p(x_0) + p(x_1) + \tilde{c}(0,\delta_{\{x_0,x_1\}}) &= \frac{k_0}{2}\Big[1 + E_0\Big((-1)^{|A\setminus\{x_0,x_1\}|}\Big)\Big] \\ &= k_0 P(1_A(x_0) \neq 1_A(x_1)), \end{aligned}$$

and so

$$P_0(1_A(x_0) \neq 1_A(x_1)) = (p(x_0) + p(x_1) + \tilde{c}(0,\delta_{\{x_0,x_1\}})) k_0^{-1}. \tag{3.5}$$

For any two distinct non-zero points, x_0 and x_1, we have

$$P_0(1_A(x_0) \neq 1_A(x_1)) = P_0(x_0 \in A) + P_0(x_1 \in A) - 2P_0(x_0 \in A, x_1 \in A).$$

Therefore, we see that (3.4) and (3.5) imply

$$\begin{aligned} P_0(\{x_0,x_1\} \subset A) &= \frac{1}{2}\Big[P_0(x_0 \in A) + P_0(x_1 \in A) - P_0(1_A(x_0) \neq 1_A(x_1))\Big] \\ &= [\tilde{c}(0,\delta_{x_0}) + \tilde{c}(0,\delta_{x_1}) - \tilde{c}(0,\delta_{\{x_0,x_1\}})](2k_0)^{-1}, \end{aligned}$$

which gives the simple bound

$$P_0(\{x_0, x_1\} \subset A) \le \frac{3}{2}\|\tilde{c}\|_\infty k_0^{-1}.$$

Note that if $0 \neq x \in A$ but $A \neq \{x\}$, then P_0-a.s. A must contain x and another non-zero point as $|A|$ is a.s. odd, and so for $0 < |x| < R_1$ and $R_2 > R_1$,

$$\begin{aligned} P_0(A = \{x\}) &\ge P_0(x \in A) - \sum_{x_1 \notin \{0,x\}} P_0(\{x, x_1\} \subset A) \\ &\ge k_0^{-1}\Big[p(x) + \tilde{c}(0, \delta_x) - (3/2)\|\tilde{c}\|_\infty (2R_2+1)^d - \sum_{|x_1|>R_2} P_0(x_1 \in A)\Big] \\ &\ge k_0^{-1}[p(x) - (1 + 2(2R_2+1)^d)\|\tilde{c}\|_\infty - \sum_{|x_1|>R_2} (p(x_1) + \tilde{c}(0, \delta_{x_1}))]. \end{aligned}$$

We have used the previous displays and (3.4) in the above. Recalling the bounds in our assumption (3.1) on $\tilde{c}$, we conclude that

$$P_0(A = \{x\}) \ge k_0^{-1}\Big[p(x) - \sum_{|x_1|>R_2} p(x_1) - 2(1 + (2R_2+1)^d)\varepsilon_2\Big]. \tag{3.6}$$

Now let

$$\eta = \eta(p(\cdot)) = \inf\{p(x) : |x| < R_1 \text{ and } p(x) > 0\} > 0,$$

choose $R_2 = R_2(p(\cdot)) > R_1$ so that $\sum_{|x_1|>R_2} p(x_1) < \eta/3$, and define

$$\varepsilon_2 = \frac{\eta}{6((2R_1+1)^d + 1)}.$$

Then by (3.6)

$$P_0(A = \{x\}) \ge (3k_0)^{-1} p(x) \text{ for all } 0 < |x| < R_1,$$

and we are done. □

For the rest of this Section we assume $\{\xi^\varepsilon : 0 < \varepsilon \le \varepsilon_0\}$ is a voter model perturbation with rate function c_ε (so that (1.10)-(1.15) hold) which is also cancellative for each ε as above with dual kernels q_0^ε satisfying (1.16)-(1.18). In particular the $\tilde{c}$ in Lemma 3.2 is now $\varepsilon^2 c_\varepsilon^*$. By Lemma 2.1, all the conclusions of that result hold.

Corollary 3.3. *Assume that*

$$\text{for small enough } \varepsilon,\ q_0^\varepsilon(A) > 0 \text{ for some } A \in Y \text{ with } |A| > 1. \tag{3.7}$$

Then there is an $\varepsilon_3 > 0$ depending on p, ε_1, $\{g_i^\varepsilon\}$ and the ε required in (3.7) so that if $0 < \varepsilon < \varepsilon_3$, then the annihilating dual with kernel q_0^ε is irreducible.

Proof. Let R_1 be as in Lemma 3.2. An easy calculation shows that

$$\|\tilde{c}\|_\infty \vee \left(\sum_x |\tilde{c}(0,\delta_x)|\right) \le \varepsilon^2[\varepsilon_1^{-2} + \vee_{i=0}^1 \|g_i^\varepsilon\|_\infty] \le \varepsilon^2 C,$$

for some constant C, independent of ε. Therefore for $\varepsilon < \varepsilon_3$ (ε_3 as claimed) we have the hypotheses, and hence conclusion, of Lemma 3.2. This allows us to apply Lemma 3.1 with $r(\cdot) = p(\cdot \,|\, |x| < R_1)$ and hence conclude that the annihilating dual ζ is irreducible for such ε. □

Remark 5. *Clearly* (3.7) *is a necessary condition for the conclusion to hold. In fact if it fails, it is easy to check that $c_\varepsilon(x,\xi)$ is a multiple of the voter model rates with random walk kernel $q_0^\varepsilon(\{x\})$. Hence this condition just eliminates voter models for which the conclusions of Corollary 3.3, as well as Lemma 2.2 and Proposition 4.1 below, will also fail in general.*

Note that if (3.7) *fails, then for some $\varepsilon_n \downarrow 0$,*

$$c^*_{\varepsilon_n}(0,\xi) = \varepsilon_n^{-2} c_{\varepsilon_n}(0,\xi) - \varepsilon_n^{-2} c_{VM}(0,\xi) = \lambda_n \tilde{c}^n_{VM}(0,\xi) - \varepsilon_n^{-2} c_{VM}(0,\xi),$$

where $\tilde{c}^n_{VM}(0,\xi)$ is the rate function for the voter model with kernel $q_0^{\varepsilon_n}(\{\cdot\})$. From this it is easy to check that if $\langle\cdot\rangle_u$ is expectation with respect to the voter model equilibrium for c_{VM} with density u, then

$$\langle (1-\xi) c^*_{\varepsilon_n}(0,\xi) - \xi c^*_{\varepsilon_n}(0,\xi)\rangle_u = 0,$$

and so by (1.14) *and* (1.23), *$f(u) \equiv 0$. Therefore, the condition $f'(0) > 0$ in Theorem 1.2 implies* (3.7).

Next we prove an irreducibility property for the voter model perturbations ξ_t^ε themselves. To do so we introduce the (unscaled) graphical representation for ξ_t^ε used in [5]. First put

$$\bar{c} = \sup_{\varepsilon<\varepsilon_0} (\|g_1^\varepsilon\|_\infty + \|g_0^\varepsilon\|_\infty + 1) < \infty.$$

For $x \in \mathbb{Z}^d$, introduce independent Poisson point processes on $\mathbb{R}_+$, $\{T_n^x, n \ge 1\}$ and $\{T_n^{*,x}, n \ge 1\}$, with rates 1 and $\varepsilon^2\bar{c}$, respectively. For $x \in \mathbb{Z}^d$ and $n \ge 1$, define independent random variables $X_{x,n}$ with distribution $p(\cdot)$, $Z_{x,n} = (Z^1_{x,n}, \ldots, Z^{N_0}_{x,n})$ with distribution $q_Z(\cdot)$, and $U_{x,n}$ uniform on $(0,1)$. These random variables are independent of the Poisson processes and all are independent of any initial condition $\xi_0^\varepsilon \in \{0,1\}^{\mathbb{Z}^d}$. For all $x \in \mathbb{Z}^d$ we allow $\xi_t^\varepsilon(x)$ to change only at times $t \in \{T_n^x, T_n^{*,x}, n \ge 1\}$. At the voter times $T_n^x, n \ge 1$ we draw a voter arrow from (x, T_n^x) to $(x + X_{x,n}, T_n^x)$ and set $\xi^\varepsilon_{T_n^x}(x) = \xi^\varepsilon_{T_n^x -}(x + X_{x,n})$. At the times $T_n^{*,x}$, $n \ge 1$ we draw "*-arrows" from $(x, T_n^{*,x})$ to each $(x + Z^i_{x,n}, T_n^{*,x})$, $1 \le i \le N_0$, and if $\xi^\varepsilon_{T_n^{*,x}-}(x) = i$ we set $\xi^\varepsilon_{T_N^{*,x}}(x) = 1 - i$ if

$$U_{x,n} < g^\varepsilon_{1-i}(\xi^\varepsilon_{T_n^{*,x}-}(x + Z^1_{x,n}), \ldots, \xi^\varepsilon_{T_n^{*,x}-}(x + Z^{N_0}_{x,n}))/\bar{c}.$$

As noted in Section 2 of [5], this recipe defines a pathwise unique process ξ_t^ε whose law is specified by the flip rates in (1.10). We refer to this as the graphical

construction of ξ_t^ε. For $x \in \mathbb{Z}^d$, $\{(X_{(x,n)}, T_n^x) : n \in \mathbb{N}\}$ and $\{(Z_{x,n}, T_n^{*,x}, U_{x,n}) : n \in \mathbb{N}\}$ are the points of independent collections of independent Poisson point processes, $(\Lambda_w^x(dy, dt), x \in \mathbb{Z}^d)$ and $(\Lambda_r^x(dy, dt, du), x \in \mathbb{Z}^d)$, on $\mathbb{Z}^d \times \mathbb{R}_+$ with rate $dt\, p(\cdot)$, and on $\mathbb{Z}^d \times \mathbb{R}_+ \times [0,1]$ with rate $\varepsilon^2 \bar{c} dt\, q_Z(\cdot)\, du$, respectively. For $R \subset \mathbb{R}^d$ and $0 \le t_1 \le t_2$ we let

$$\mathcal{G}(R \times [t_1, t_2]) = \sigma\Big(\Lambda_w^x|_{\mathbb{Z}^d \times [t_1,t_2]}, \Lambda_r^{x'}|_{\mathbb{Z}^d \times [t_1,t_2] \times [0,1]} : x, x' \in R\Big),$$

that is, the σ-field generated by the points of the graphical construction in $R \times [t_1, t_2]$.

A coalescing branching random walk dual for ξ_t^ε is constructed in [5]. We give here only the part of that dual which we need. Using only the Poisson processes $T_n^x, x \in \mathbb{Z}^d$, define a coalescing random walk system as follows. Fix $t > 0$. For each $y \in \mathbb{Z}^d$ define $B_u^{y,t}, u \in [0,t]$ by putting $B_0^{y,t} = y$ and then proceeding "down" in the graphical construction and using the voter arrows to jump. More precisely, if $T_1^y > t$ put $B_u^{y,t} = y$ for all $u \in [0,t]$. Otherwise, choose the largest $T_j^y = s < t$, and put $B_u^{y,t} = y$ for $u \in [0, t-s)$ and $B_{t-s}^{y,t} = x + X_{x,j}$. Continue in this fashion to complete the construction of $B_u^y, u \in [0,t]$. Note that each $B_u^{y,t}$ is a rate one random walk with step distribution $p(\cdot)$ and that the walks coalesce when they meet: if $B_u^{x,t} = B_u^{y,t}$ for some $u \in [0,t]$ then $B_s^{x,t} = B_s^{y,t}$ for all $u \le s \le t$. On the event that no $*$-arrow is encountered along the path $B_\cdot^{x,t}$, i.e., $(z, T_n^{*,z}) \ne (B_{t-u}^{x,t}, t-u)$ for all z, n and $0 \le u \le t$, then

$$\xi_t^\varepsilon(x) = \xi_0^\varepsilon(B_t^{x,t}) \ \forall\ \xi_0^\varepsilon \in \{0,1\}^{\mathbb{Z}^d}. \tag{3.8}$$

Lemma 3.4. *Fix $t > 0$, distinct $y_0, y_1 \in \mathbb{Z}^d$ and finite disjoint $B_0, B_1 \subset \mathbb{Z}^d$. Then there exists a finite $\Lambda = \Lambda(y_0, y_1, B_0, B_1) \subset \mathbb{Z}^d$ and a $\mathcal{G}(\Lambda \times [0,t])$-measurable event $G = G(t, y_0, y_1, B_0, B_1)$ such that $P(G) > 0$ and on G,*

(i) $T_1^{*,z} > t$ for all $z \in \Lambda$,

(ii) $B_u^{x,t} \in \Lambda$ for all $x \in B_0 \cup B_1$, $u \in [0,t]$,

(iii) $B_t^{x,t} = y_i$ for all $x \in B_i$, $i = 0, 1$.

If $\xi_0^\varepsilon(y_i) = i$, $i = 0, 1$, then on the event G, $\xi_t^\varepsilon(x) = i$ for all $x \in B_i$, $i = 0, 1$.

Proof. We reason as in the proof of Lemma 3.1, but now working with the dual of ξ^ε, using the fact that the $B_u^{y,t}$ are independent, irreducible random walks as long as they don't meet. There are sets B_0', B_1' which are far apart, each with widely separated points, such that a sequence of walk steps can move the walks from B_0 to B_0' and B_1 to B_1' without collisions. If B_0' and B_1' are sufficiently far apart, then by irreducibility there is a sequence of steps resulting in the walks from B_0' coalescing at some site y_0', the walks from B_1' coalescing at some site y_1', all without collisions between the two collections of walks, and with y_0' and y_1' far apart. Now by moving one walk at a time it is possible to prescribe a set of walk steps which take the two walks from y_0 and y_1 to y_0' and y_1', respectively, without collisions between the two walks. By reversing these steps (recall p is symmetric) we can therefore have the above walks follow steps which will take them from y_0' and y_1' to y_0 and y_1, respectively, without collisions. In this way we can prescribe walk steps which occur with positive probability and ensure

that $B_t^{x,t} = y_i$ for all $x \in B_i$. Let Λ be a finite set large enough to contain all the positions of the walks in this process, and let G be the event that $T_1^{*,x} > t$ for all $x \in \Lambda$, and such that the T_n^x and $X_{x,n}$, $x \in \Lambda$, allow for the above prescribed sequence of walk steps to occur by time t. Then G has the desired properties, and on this event, $\xi_t^\varepsilon(x) = \xi_0^\varepsilon(B_t^{x,t})$ for all $x \in B_0 \cup B_1$ by (3.8). Now the fact that (iii) holds on G, implies the final conclusion by the choice of y_i. □

In addition to Lemma 3.4 we will need the simpler fact that for any fixed $t > 0$ and $z \in \mathbb{Z}^d$,

$$\inf_{\xi_0^\varepsilon : \xi_0^\varepsilon(0)=1} P(\xi_t^\varepsilon(z) = 1) > 0. \tag{3.9}$$

This is clear because there is a sequence of random walk steps leading from 0 to z and there is positive probability that the walk makes these steps before time t and that no other transitions occur at any site in the sequence.

Remark 6. *It is clear that the above holds equally well for voter model perturbations in $d = 2$.*

4. A complete convergence theorem for cancellative systems

To make effective use of annihilating duality we will need to know that for large t, if $\xi_t \neq \mathbf{0}, \mathbf{1}$ and finite $A \subset \mathbb{Z}^d$ is large, then there will be many sites in $\xi_t \cap A$ which can flip values in a fixed time interval, and that the probability there will be an odd number of these flips is close to 1/2. For $x \in \mathbb{Z}^d$ and $A \subset \mathbb{Z}^d$ define

$$A(x, \xi) = \{y \in A : \xi(y) = 1 \text{ and } \xi(y + x) = 0\}.$$

The conditions we will use are: there exists $x_0 \in \mathbb{Z}^d$ such that

$$\lim_{K\to\infty} \sup_{\substack{A\subset\mathbb{Z}^d \\ |A|\geq K}} \limsup_{t\to\infty} P\Big(|\xi_t| > 0 \text{ and } A(x_0, \xi_t) = \emptyset\Big) = 0 \text{ if } |\widehat{\xi}_0| = \infty, \tag{4.1}$$

and

$$\lim_{K\to\infty} \sup_{\substack{A\in Y, \xi_0\in\{0,1\}^{\mathbb{Z}^d}: \\ |A(x_0,\xi_0)|\geq K}} \Big|P(|\xi_1 \cap A| \text{ is odd}) - \tfrac{1}{2}\Big| = 0 \tag{4.2}$$

We will verify in Lemmas 4.2 and 5.3 below that our voter model perturbations have these properties for all sufficiently small ε, but first we will show how they are used along with (2.7) to obtain complete convergence of ξ_t. Recall $\nu_{1/2}$ is the translation invariant stationary measure in (1.22).

Proposition 4.1. *Let ξ_t be a translation invariant cancellative spin-flip system with rate function $c(x, \xi)$ satisfying (1.15)-(1.18), (4.1), and (4.2). Let ζ_t be the annihilating dual with Q-matrix given in (1.19) and assume that (2.7) holds. Then $\nu_{1/2}$ satisfies (1.4), and if $|\widehat{\xi}_0| = \infty$ then*

$$\xi_t \Rightarrow \beta_0(\xi_0)\delta_0 + (1 - \beta_0(\xi_0))\nu_{1/2} \text{ as } t \to \infty. \tag{4.3}$$

Proof. We start with some preliminary facts. First, (4.1) implies that for any $m < \infty$ and $\xi_0 \in \{0,1\}^{\mathbb{Z}^d}$ with $|\widehat{\xi_0}| = \infty$,

$$\lim_{K\to\infty} \sup_{\substack{A\subset\mathbb{Z}^d \\ |A|\geq K}} \limsup_{t\to\infty} P\Big(|\xi_t| > 0 \text{ and } |A(x_0,\xi_t)| < m\Big) = 0. \tag{4.4}$$

This is because $|A| \geq mK$ implies A can be written as the disjoint union of sets $A_1, \ldots, A_m$ with each $|A_i| \geq K$, and

$$\{|\xi_t| > 0 \text{ and } |A(x_0,\xi_t)| < m\} \subset \cup_{i=1}^m \{|\xi_t| > 0 \text{ and } |A_i(x_0,\xi_t)| = 0\}.$$

Applying (4.1) we obtain (4.4).

Next, we need a slight upgrade of the basic duality equation. As shown in [20], (2.1) can be extended by applying the Markov property of ξ_t at a time $v < t$. If the processes ξ_t and ζ_t are independent, then for all $u, v \geq 0$,

$$P(|\xi_{v+u} \cap \zeta_0| \text{ is odd }) = P(|\xi_v \cap \zeta_u| \text{ is odd }). \tag{4.5}$$

Let $\nu_{1/2}$ be defined by (2.4). Since we are assuming $|\widehat{\xi_0}| = \infty$, we have $\beta_1(\xi_0) = 0$ by (1.8). In view of (2.4) and $\delta_{\mathbf{0}}(|\xi \cap A| \text{ is odd}) = 0$, to prove (4.3) it suffices to prove (recall (2.3)) that for fixed $A \in Y$,

$$\lim_{t\to\infty} P\Big(|\xi_t \cap A| \text{ is odd}\Big) = \tfrac{1}{2}\beta_\infty(\xi_0) P(\zeta_t^A \neq \emptyset \ \forall\ t \geq 0). \tag{4.6}$$

Fix $\varepsilon > 0$. By (4.2) there exists $K_1 < \infty$ such that if $B \in Y$ and $|B(x_0,\xi_0)| \geq K_1$ then

$$\Big|P(|\xi_1 \cap B| \text{ is odd}) - \frac{1}{2}\Big| < \varepsilon. \tag{4.7}$$

By (4.4), there exists $K_2 < \infty$ and $s_0 < \infty$ such that if $|B| \geq K_2$ and $s \geq s_0$ then

$$P(\xi_s \neq \emptyset \text{ and } |B(x_0,\xi_s)| < K_1) < \varepsilon. \tag{4.8}$$

By (2.7) we can choose $T = T(A, K_2) < \infty$ large enough so that

$$P(0 < |\zeta_T^A| \leq K_2) < \varepsilon. \tag{4.9}$$

For $t > 1+T+s_0$ let $s = t-(1+T)$ and put $u = T$ and $v = s+1$ in (4.5). Then $P(|\xi_t \cap A| \text{ is odd}) = P(|\xi_{s+1} \cap \zeta_T^A| \text{ is odd})$, where ξ_t and ζ_t^A are independent. (At this point the reader may want to consult Figure 1 and Remark 7 below.) Making use of the Markov property of ξ_t we obtain

$$\begin{aligned} &P(|\xi_t \cap A| \text{ is odd}) - \tfrac{1}{2}P(\xi_s \neq \emptyset)P(\zeta_T^A \neq \emptyset) \\ &\qquad = \sum_{B\neq\emptyset} P(\zeta_T^A = B)\Big[P(|\xi_{s+1} \cap B| \text{ is odd}) - \tfrac{1}{2}P(\xi_s \neq \emptyset)\Big] \\ &\qquad = \sum_{B\neq\emptyset} P(\zeta_T^A = B) E\Big[(E_{\xi_s}(|\xi_1 \cap B| \text{ is odd}) - \tfrac{1}{2})\, 1\{\xi_s \neq \emptyset\}\Big]. \end{aligned}$$

By (4.9),

$$(4.10)\quad |P(|\xi_t \cap A| \text{ is odd}) - \tfrac{1}{2}P_\xi(\xi_s \neq \emptyset)P(\zeta_T^A \neq \emptyset)| < \varepsilon + \sum_{|B|>K_2} P(\zeta_T^A = B)E\Big[\Big|E_{\xi_s}(|\xi_1 \cap B| \text{ is odd}) - \tfrac{1}{2}\Big| 1\{\xi_s \neq \emptyset\}\Big].$$

By (4.8), since $s > s_0$, each expectation in the last sum is bounded above by

$$(4.11)\qquad \varepsilon + E\Big[\Big|E_{\xi_s}(|\xi_1 \cap B| \text{ is odd}) - \tfrac{1}{2})\Big| 1\{B(x_0, \xi_s) \geq K_1\}\Big].$$

Applying the bound (4.7) in this last expression, and then combining (4.10) and (4.11) we obtain

$$|P(|\xi_t \cap A| \text{ is odd}) - \tfrac{1}{2}P(\xi_s \neq \emptyset)P(\zeta_T^A \neq \emptyset)| < 3\varepsilon.$$

Let t (and hence s) tend to infinity, and then T tend to infinity above to complete the proof of (4.6), and hence (4.3).

Finally, let $|\xi_0| = |\widehat{\xi}_0| = \infty$. Then $\beta_0(\xi_0) = 0$ by (1.8), so (4.3) and (4.4) imply that for any finite m, $\nu_{1/2}(|\xi| \geq m, |\widehat{\xi}| \geq m) = 1$, and this implies coexistence. □

Remark 7. *Figure 1 below gives a graphical view of the above argument. Time runs up for ξ and down for ζ. Conditional on $\xi_s \neq \emptyset$ and $\zeta_T^A \neq \emptyset$, (2.7) guarantees that $B = \zeta_T^A$ is large and (4.5) guarantees $B(x_0, \xi_s)$ is large. In the dashed boxes in Figure 1, the $(\bullet\ \circ)$ pairs indicate the locations $(x, x+x_0)$, $x \in B(x_0, \xi_s)$. Finally, (4.2) now guarantees that $|\xi_{s+1} \cap B|$ will be odd with probability approximately $\frac{1}{2}$.*

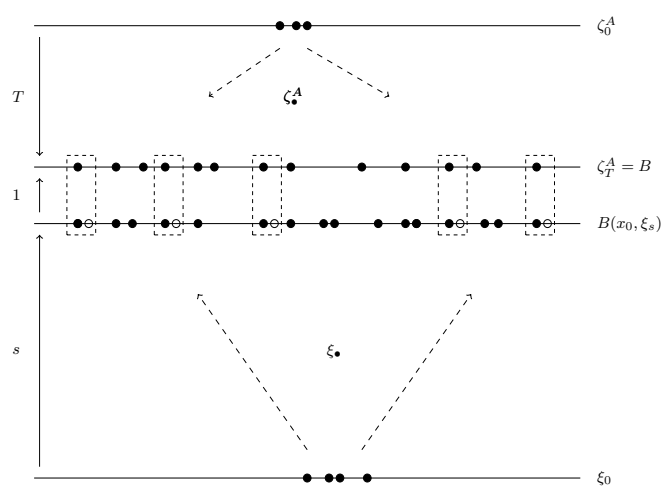

Fig 1: $P(|\xi_{s+1} \cap \zeta_T^A| \text{ is odd}) \approx \frac{1}{2}P(\xi_s \neq \emptyset)P(\zeta_T^A \neq \emptyset)$

Verification of (4.1) for our voter model perturbations requires a comparison with oriented percolation which we will save for the next section. Here we present a proof of (4.2), based heavily on ideas from [1]. See Lemma 7 of [28] for a purely cancellative version of this result.

Lemma 4.2. *If $\xi^\varepsilon_\cdot$ is a voter model perturbation, then there exists $\varepsilon_1 > 0$ and $x_0 \in \mathbb{Z}^d$ such that (4.2) holds for $\xi^\varepsilon_\cdot$ if $\varepsilon < \varepsilon_1$.*

Proof. Fix any x_0 with $p(x_0) > 0$. We will prove that if $\delta > 0$ then there exists K such that if $|A(x_0, \xi_0)| \geq K$ then

$$\left| P(|\xi^\varepsilon_1 \cap A| \text{ is odd}) - \tfrac{1}{2} \right| < \delta. \tag{4.12}$$

Using the graphical construction of ξ^ε_t described in Section 3, we define a version of the "almost isolated sites" of [1]. First we give the informal definition. For $x \in \mathbb{Z}^d$, let $U(x)$ be the indicator of the event that during the time period $[0, 1]$, no change can occur at site $x + x_0$ and no change can occur at x except possibly due to a (first) voter arrow directed from x to $x + x_0$. Let $V(x)$ be the indicator of the event that during the time period $[0, 1]$ no site y outside $\{x, x + x_0\}$ can change due to the value at x. More formally, for $y \in \mathbb{Z}^d$ and $A \in Y$ with $|A| \leq N_0$, define

$$\begin{aligned}\tau(y, A) = \min\{T^y_n : A = \{X_{y,n}\}, n \in \mathbb{N}\} \\ \wedge \min\{T^{*,y}_n : A = \{Z^1_{y,n}, \ldots, Z^{N_0}_{y,n}\}, n \in \mathbb{N}\},\end{aligned}$$

and $\tau(y) = \min\{\tau(y, A) : A \in Y\}$. We can now define

$$\begin{aligned}U(x) &= 1\{\tau(x + x_0) > 1, X_{x,1} = x_0, T^x_2 > 1 \text{ and } T^{*,x}_1 > 1\}, \text{ and} \\ V(x) &= 1\{\tau(y, A) > 1 \ \forall\, y \in \mathbb{Z}^d \setminus \{x, x + x_0\} \text{ and } A \in Y : x \in y + A\},\end{aligned}$$

and call x almost isolated if $U(x)V(x) = 1$.

By standard properties of Poisson processes,

$$\tau(x, A) \text{ and } \tau(y, B) \text{ are independent whenever } x \neq y \text{ or } A \neq B. \tag{4.13}$$

We also define

$$\nu(A) = P(\{X_{0,1}\} = A) + P(\{Z^1_{0,1}, \ldots, Z^{N_0}_{0,1}\} = A),$$

and observe that $\nu(A) = 0$ if $|A| > N_0$, and $\sum_{A \in Y} \nu(A) = 2$. Use the fact that $\{T^0_n : \{X_{0,n}\} = A\}$ are the points of a Poisson point process with rate $P(\{X_{0,1}\} = A)$, and $\{T^{*,0}_n : \{Z^1_{0,n}, \ldots, Z^{N_0}_{0,n}\} = A\}$ are the points of an independent Poisson point process with rate $\bar{c}\varepsilon^2 P(A = \{Z^1_{0,n}, \ldots, Z^{N_0}_{0,n}\})$ to conclude that

$$\begin{aligned}P(\tau(y, A) > 1) &= \exp\Big(-P(\{X_{0,1}\} = A) - \bar{c}\varepsilon^2 P(\{Z^1_{0,n}, \ldots, Z^{N_0}_{0,n}\} = A)\Big) \\ &\geq e^{-(1+\bar{c})\nu(A)}.\end{aligned} \tag{4.14}$$

For each $x \in \mathbb{Z}^d$, the variables $U(x), V(x)$ are independent (this much is clear from (4.13)), and we claim that $u_0 = E(U(x))$ and $v_0 = E(V(x))$ are positive uniformly in ε. To check this for v_0 we apply (4.14) to get

$$\begin{aligned} v_0 &\geq \exp\Big(-(1+\bar{c}) \sum_{A \in Y} \sum_{y \in \mathbb{Z}^d \setminus \{x\}} \nu(A) 1\{x - y \in A\} \Big) \\ &\geq \exp\Big(-(1+\bar{c}) \sum_{A \in Y} |A| \nu(A) \Big) \end{aligned}$$

which is positive, uniformly in $\varepsilon \leq \varepsilon_0$. For u_0, we have by the choice of x_0,

$$u_0 \geq \exp\Big\{ -(1+\bar{c}) \sum_{A \in Y} \nu(A) \Big\} p(x_0) P(T_2^0 > 1) P(T_1^{*,0} > 1) > 0.$$

If $w_0 = w_0(\varepsilon) = u_0 v_0$, then we have verified that

$$\gamma = \min\{w_0(\varepsilon) : 0 < \varepsilon \leq \varepsilon_0\} > 0. \tag{4.15}$$

Now suppose $\mathcal{Y} = \{y_1, \ldots, y_J\} \subset \mathbb{Z}^d$ and $|y_i - y_j| > 2|x_0|$ for $i \neq j$. Then $U(y_1), \ldots, U(y_J)$ are independent but $V(y_1), \ldots, V(y_J)$ are not. Nevertheless, we claim they are almost independent if all $|y_i - y_j|$, $i \neq j$, are large, and hence if we let $W(y_i) = U(y_i)V(y_i)$, then $W(y_1), \ldots, W(y_J)$ are almost independent. More precisely, we claim that for any $J \geq 2$ and $a_i \in \{0, 1\}$, $1 \leq i \leq J$,

$$\lim_{n \to \infty} \sup_{\substack{\mathcal{Y} = \{y_1, \ldots, y_J\}, \\ |y_i - y_j| \geq n \, \forall \, i \neq j}} \Big| P(W(y_i) = a_i \; \forall 1 \leq i \leq J) - \prod_{i=1}^{J} P(W(y_i) = a_i) \Big| = 0 \tag{4.16}$$

For the time being, let us suppose this fact.

Given J and $\mathcal{Y} = \{y_1, \ldots, y_J\}$ let $S(J, \mathcal{Y}) = \sum_{y \in \mathcal{Y}} W(y)$. Then (4.16) implies $S(J, \mathcal{Y})$ is approximately binomial if the y_i are well separated. That is, if $\mathcal{B}(J, w_0)$ is a binomial random variable with parameters J and w_0, and

$$\Delta(J, \mathcal{Y}, k) = \Big| P(S(J, \mathcal{Y}) = k) - P(\mathcal{B}(J, w_0) = k) \Big|$$

then (4.16) implies that for $k = 0, \ldots, J$,

$$\lim_{n \to \infty} \sup_{\substack{\mathcal{Y} = (y_1, \ldots, y_J) \\ |y_i - y_j| \geq n, i \neq j}} \Delta(J, \mathcal{Y}, k) = 0. \tag{4.17}$$

Now fix $\delta > 0$. A short calculation shows that

$$p_0 = P(T_1^x > 1 | U(x) = 1) = P(T_1^x > 1 | T_2^x > 1) = \frac{1}{2}. \tag{4.18}$$

By (4.15) and (4.17), we may choose $J = J(\delta)$ such that

$$(1 - \gamma)^J < \delta, \tag{4.19}$$

and then $n = n(J, \delta)$ so that for all $\mathcal{Y} = \{y_1, \ldots, y_J\}$ with $|y_i - y_j| \geq n$ for $i \neq j$,

$$\Delta(J, \mathcal{Y}, 0) < \delta. \tag{4.20}$$

Given J and n, it is easy to see that there exists $K = K(J, n)$ such that if $B \subset \mathbb{Z}^d$ and $|B| \geq K$ then B must contain some $\mathcal{Y} = \{y_1, \ldots, y_J\}$ such that $|y_i - y_j| \geq n$ for $i \neq j$.

Now suppose that $|A(x_0, \xi_0^\varepsilon)| \geq K$ and $\mathcal{Y} = \{y_1, \ldots, y_J\} \subset A(x_0, \xi_0^\varepsilon)$ with $|y_i - y_j| \geq n$ for all $i \neq j$. Let $\mathcal{I}$ be the set of y_j with $W(y_j) = 1$, so that $|\mathcal{I}| = S(J, \mathcal{Y})$.

Let $\mathcal{G}$ be the σ-field generated by

$$\begin{aligned} &\{1(y_j \in \mathcal{I}) : j = 1, \ldots, J\} \\ &\qquad \cup \{1\{x \in \mathcal{I}^c\}(T_n^x, T_n^{*,x}, X_{x,n}, Z_{x,n}, U_{x,n} : x \in \mathbb{Z}^d, n \geq 1)\}. \end{aligned} \tag{4.21}$$

If $g_j = 1\{T_1^{y_j} > 1\}$, then conditional on $\mathcal{G}$,

$$\{g_j : y_j \in \mathcal{I}\} \text{ are iid Bernoulli rv's with mean } p_0 = \tfrac{1}{2} \tag{4.22}$$

and $X = \sum_j 1\{y_j \in \mathcal{I}\} g_j$ is binomial with parameters $(|\mathcal{I}|, p_0 = \frac{1}{2})$. This is easily checked by conditioning on the $\mathcal{G}$-measurable set $\mathcal{I}$ and using (4.18).

Let $h = \sum_{x \in A} 1\{x \notin \mathcal{I}) \xi_1^\varepsilon(x)$. Then at time 1 we have the decomposition

$$|\xi_1^\varepsilon \cap A| = h + X, \tag{4.23}$$

where we have used the fact that for $y_j \in \mathcal{I}$, $\xi_s^\varepsilon(y_j)$ will flip from a $1 = \xi_0^\varepsilon(y_j)$ to a $0 = \xi_0^\varepsilon(y_j + x_0)$ during the time interval $[0, 1]$ iff $g_j = 0$. Since h is $\mathcal{G}$-measurable,

$$P(|\xi_1^\varepsilon \cap A| \text{ is odd} \mid \mathcal{G})(\omega) = P(X = 1 - h(\omega) \bmod 2 \mid \mathcal{G})(\omega) = \frac{1}{2} \text{ a.s. on } \{|\mathcal{I}| > 0\},$$

the last by an elementary binomial calculation and the fact that conditional on $\mathcal{G}$, X is binomial with parameters $(|\mathcal{I}|, \frac{1}{2})$. Take expectations in the above and use (4.20) and then (4.19) to conclude that

$$\begin{aligned} |P(|\xi_1^\varepsilon \cap A| \text{ is odd}) - \tfrac{1}{2}| &\leq P(|\mathcal{I}| = 0) \\ &\leq \delta + P(\mathcal{B}(J, \gamma) = 0) \\ &\leq \delta + (1 - \gamma)^J \leq 2\delta. \end{aligned} \tag{4.24}$$

This yields inequality (4.12).

It remains to verify (4.16). This is easy if p and q_Z have finite support (see Remark 8 below). In general, the idea is to write

$$V(y_i) = V_1(y_i) V_2(y_i) V_3(y_i) \tag{4.25}$$

where $V_1(y_1), \ldots, V_1(y_J)$ are independent and independent of $U(y_1), \ldots, U(y_J)$, and $V_2(y_1), V_3(y_1), \ldots, V_2(y_J), V_3(y_J)$ are all one with high probability if the

y_i are sufficiently spread out. Let $n \geq 2|x_0|$ and $\mathcal{Y} = \{y_1, \ldots, y_J\}$ be given with $|y_i - y_j| \geq n$, and let $\mathcal{Y}_0 = \{y_1 + x_0, \ldots, y_J + x_0\}$. Note that $y_1, y_1 + x_0, \ldots, y_J, y_J + x_0$ are distinct. Define

$$\begin{aligned}
V_1(y_i) &= 1\{\tau(z, A) > 1 \ \forall \ z \notin \mathcal{Y} \cup \mathcal{Y}_0, A \in Y : (z + A) \cap \mathcal{Y} = \{y_i\}\} \\
V_2(y_i) &= 1\{\tau(z, A) > 1 \ \forall \ z \notin \mathcal{Y} \cup \mathcal{Y}_0, A \in Y : z + A \supset \{y_i, y_j\} \\
&\qquad \text{for some } j \neq i\}, \\
V_3(y_i) &= 1\{\tau(z, A) > 1 \ \forall \ z \in (\mathcal{Y} \cup \mathcal{Y}_0) \setminus \{y_i, y_i + x_0\}, A \in Y : y_i \in z + A\}.
\end{aligned}$$

A bit of elementary logic shows that (4.25) holds. If a pair (z, A) occurs in the definition of some $V_1(y_i)$ then it cannot occur in any $V_1(y_j), j \neq i$, and hence $V_1(y_1), \ldots, V_1(y_J)$ are independent, and also independent of $U(y_1), \ldots, U(y_J)$. Therefore, to prove (4.16) it suffices to prove that

$$\lim_{n\to\infty} \sup_{\substack{\mathcal{Y}=\{y_1,\ldots,y_J\}, \\ |y_i - y_j| \geq n \ \forall \ i \neq j}} P(V_2(y_i)V_3(y_i) \neq 1) = 0. \tag{4.26}$$

We treat $V_3(y_i)$ first. By (4.14),

$$\begin{aligned}
P(V_3(y_i) = 1) &\geq \exp\Big(- (1 + \bar{c}) \sum_{z \in (\mathcal{Y} \cup \mathcal{Y}_0) \setminus \{y_i, y_i + x_0\}} \sum_{A \in Y} \nu(A) 1\{y_i \in z + A\}\Big) \\
&\geq \exp\Big(- 2J(1 + \bar{c}) \sum_{A \in Y} \nu(A) 1\{\text{diam}(A) > n\}\Big) \\
&\to 1 \text{ as } n \to \infty.
\end{aligned}$$

To treat $V_2(y_i)$ we note that if a pair (z, A) occurs in the definition of $V_2(y)$, then $\text{diam}(A) \geq n$, so

$$\begin{aligned}
&P(V_2(y_i) = 1) \\
&\quad \geq \exp\Big(- (1 + \bar{c}) \sum_{B \subset \mathcal{Y}} \sum_{A \in Y} \sum_{z \neq y_i} 1\{y_i \in B = (z + A) \cap \mathcal{Y}, \text{diam}(A) \geq n\} \nu(A)\Big).
\end{aligned}$$

In the sum above, given $B \ni y_i$ there at most $|A|$ choices for z such that $(z + A) \cap \mathcal{Y} = B$. In fact, there are at most $|A|$ choices of z such that $y_i \in z + A$ as this implies $z \in y_i - A$. Thus

$$\begin{aligned}
P(V_2(y_i) = 1) &\geq \exp\Big(- (1 + \bar{c}) \sum_{B \subset \mathcal{Y}} \sum_{A \in Y} 1\{\text{diam}(A) \geq n\} |A| \nu(A)\Big) \\
&\geq \exp\Big(- (1 + \bar{c}) 2^J \sum_{A \in Y} 1\{\text{diam}(A) \geq n\} |A| \nu(A)\Big) \\
&\to 1 \text{ as } n \to \infty.
\end{aligned}$$

This proves (4.26) and hence (4.16). □

Remark 8. *Note that if $p(\cdot)$ and $q_Z(\cdot)$ have finite support then the proof simplifies somewhat because for large enough n the left-hand side of (4.16) is zero. This is because the A's arising in the definition of $V(x)$ will have uniformly bounded diameter which will show that for $|y_i - y_j|$ large, $V(y_1), \ldots, V(y_J)$ will be independent.*

5. Proof of Theorem 1.2

To prove Theorem 1.2 it will suffice, in view of Proposition 4.1, Lemmas 2.2 and 4.2, and Remark 4, to prove that for small enough ε, both of the conditions (2.8) and (4.1) hold for ξ_t^ε. The proof of (4.1) is given in Lemma 5.3 below after first developing the necessary oriented percolation machinery. With (4.1) in hand the proof of (2.8) is then straightforward.

We suppose now that ξ^ε is a voter model perturbation with rate function $c_\varepsilon(x,\xi)$ and that all the assumptions of Theorem 1.2 are in force. We also assume that ξ^ε is constructed using the Poisson processes $T_n^x, T_n^{*,x}$ and the variables $X_{x,n}, Z_{x,n}^i, U_{x,n}$ as in Section 3. We assume $|\hat{\xi}_0^\varepsilon| = \infty$, so that by (1.8) $\beta_1(\xi_0^\varepsilon) = 0$. By the results of [5] for small ε we expect that when ξ_t^ε survives there will be blocks in space-time, in the graphical construction, containing both 0's and 1's, which dominate a super-critical oriented percolation. The percolation process necessarily spreads out. So if $A \subset \mathbb{Z}^d$ is large, eventually there will be many blocks containing 0's and 1's near the sites of A at times just before t, allowing for many independent tries to force $|A(\xi_t, x_0)| \geq 1$.

Let $\mathbb{Z}_e^d$ be the set of $x \in \mathbb{Z}^d$ such that $\sum_i z_i$ is even. Let $\mathcal{L} = \{(x,n) \subset \mathbb{Z}^d \times \mathbb{Z}^+ : \sum_i x_i + n \text{ is even}\}$. We equip $\mathcal{L}$ with edges from (x,n) to $(x+e, n+1)$ and $(x-e, n+1)$ for all $e \in \{e_1, \ldots, e_d\}$, where e_i is the ith unit basis vector. Given a family of Bernoulli random variables $\theta(x,n), (x,n) \in \mathcal{L}$, we define open paths in $\mathcal{L}$ using the $\theta(y,n)$ and the edges in $\mathcal{L}$ in the usual way. That is, a sequence of points $z_0, \ldots, z_n$ in $\mathcal{L}$ is an open path from z_0 to z_n iff there is an edge from z_i to z_{i+1} and $\theta(z_i) = 1$ (in which case we say site z_i is open) for $i = 0, \ldots, n-1$. We will write $(x,n) \to (y,m)$ to indicate there is an open path in $\mathcal{L}$ from (x,n) to (y,m). Define the open cluster starting at $(x,n) \in \mathcal{L}$,

$$\mathcal{C}(x,n) = \big\{(y,m) \in \mathcal{L} : m \geq n \text{ and } (x,n) \to (y,m) \text{ in } \mathcal{L}\big\}.$$

For $(x,n) \in \mathcal{L}$ let $W_m^{(x,n)} = \{y : (x,n) \to (y,m)\}$, $m \geq n$. We will write W_n^0 for $W_n^{(0,0)}$. For $k = 1, \ldots, d$, say that $(x,n) \to_k (y,m)$ if there is an open path from (x,n) to (y,m) using only edges of the form $(x,n) \to (x+e_k, n+1)$ or $(x,n) \to (x-e_k, n+1)$. We define the corresponding "slab" clusters $\mathcal{C}_k(x,n)$ and processes $W_{k,m}^{(x,n)}$ using these paths. Clearly $\mathcal{C}_k(x,n) \subset \mathcal{C}(x,n)$ and $W_{k,m}^{(x,n)} \subset W_m^{(x,n)}$. If $W_0 \subset \mathbb{Z}^d$, let $W_m = \cup_{x \in W_0} W_m^{(x,0)}$.

Lemma 5.1. *Suppose the $\{\theta(z,n)\}$ are iid, and $1 - \gamma = P(\theta(x,n) = 1) \geq 1 - 6^{-4}$. Then*

$$\rho_\infty = P(|\mathcal{C}_1(0,0)| = \infty) > 0, \tag{5.1}$$

and

$$\lim_{K\to\infty} \sup_{\substack{A\subset 2\mathbb{Z}^d \\ |A|\geq K}} \limsup_{n\to\infty} P(W^0_{2n} \neq \emptyset \text{ and } W^0_{2n} \cap A = \emptyset) = 0. \tag{5.2}$$

Proof. For (5.1) see Theorem A.1 (with $M = 0$) in [11]. The limit (5.2) is known for $d = 1$, while the $d > 1$ case is an immediate consequence of the "shape theorem" for W^0_{2n}, the discrete time analogue of the shape theorem for the contact process in [14]. Since this discrete time result does not appear in the literature, we will give a direct proof of (5.2), but for the sake of simplicity will restrict ourselves to the $d = 2$ case. We need the following $d = 1$ results, which we state using our "slab" notation:

$$\exists\ \rho_1 > 0 \text{ such that } \liminf_{n\to\infty} P((x,0) \in W^0_{1,2n}) \geq \rho_1 \text{ for all } x \in 2\mathbb{Z}, \tag{5.3}$$

and for fixed $K_0 \in \mathbb{N}$,

$$\lim_{K\to\infty} \sup_{A\subset 2\mathbb{Z}\times\{0\},|A|\geq K} \limsup_{n\to\infty} P(W^0_{1,2n} \neq \emptyset, |W^0_{1,2n} \cap A)| < K_0) = 0. \tag{5.4}$$

These facts are easily derived using the methods in [10] (see also Lemma 3.5 in [16], the Appendix in [15] and Section 2 of [2]).

The idea of the proof of the $d = 2$ case of (5.2) is the following. If n is large then on the event $W^0_{2n} \neq \emptyset$ we can find, with high probability, a point $z \in W^0_{2k}$ for some small k such that $W^{(z,2k)}_{1,2m} \neq \emptyset$ for some large $m < n$. With high probability $W^{(z,2k)}_{1,2m}$ will contain many points z' from which we can start independent "e_2" slab processes $W^{(z',2m)}_{2,2n}$. Many of these will be large, providing many independent chances for $W^{(z',2m)}_{2,2n} \cap A \neq \emptyset$, forcing $W^0_{2n} \cap A \neq \emptyset$.

Here are the details. We may assume without loss of generality that all sets A considered here are finite. Fix $\delta > 0$, and choose positive integers J_0, K_0 satisfying satisfy $(1-\rho_\infty)^{J_0} < \delta$ and $(1-\rho_1)^{K_0} < \delta$. By (5.4) we can choose a positive integer K_1 such that for all $A \subset 2\mathbb{Z} \times \{0\}$, $|A| \geq K_1$,

$$\limsup_{n\to\infty} P(W^0_{1,2n} \neq \emptyset, |W^0_{1,2n} \cap A| < K_0) < \delta. \tag{5.5}$$

For $x = (x_1, x_2) \in \mathbb{Z}^2$ and $A \subset \mathbb{Z}^2$ let $\pi_1 x = (x_1, 0)$, $\pi_2 x = (0, x_2)$, and $\pi_i A = \{\pi_i a : a \in A\}$, $i = 1, 2$. Observe that at least one of the $|\pi_i A| \geq \sqrt{|A|}$. We now fix any $A \subset 2\mathbb{Z}^2$ with $|A| \geq K_1^2$, and suppose $|\pi_1 A| \geq K_1$. For convenience later in the argument, fix any $A' \subset A$ such that $\pi_1 A' = \pi_1 A$ and π_1 is one-to-one on A'. By (5.5) we may choose a positive integer $n_1 = n_1(A)$ such that

$$P(W^0_{1,2n} \neq \emptyset, |W^0_{1,2n} \cap \pi_1 A'| < K_0) < \delta \text{ for all } n \geq n_1. \tag{5.6}$$

We may increase n_1 if necessary so that $P(|\mathcal{C}_1(0,0)| < \infty, W^0_{1,2n_1} \neq \emptyset) < \delta$, which implies that

$$P(W^0_{1,2n_1} \neq \emptyset, W^0_{1,2n} = \emptyset) < \delta \text{ for all } n \geq n_1. \tag{5.7}$$

Let $m(j) = 2(j-1)n_1, j = 1, 2, \ldots$, and define a random sequence of points $z_1, z_2, \ldots$ as follows. If $W^0_{m(j)} \neq \emptyset$ let z_j be the point in $W^0_{m(j)}$ closest to the origin, with some convention in the case of ties. If $W^0_{m(j)} = \emptyset$ put $z_j = 0$. Define

$$N = \inf\{j : z_j \in W^0_{m(j)} \text{ and } W^{(z_j,m(j))}_{1,m(j+1)} \neq \emptyset\}.$$

Since $P(W^0_{1,2n_1} = \emptyset) \leq 1 - \rho_\infty$, the Markov property implies,

$$\begin{aligned} P(W^0_{m(j)} \neq \emptyset, N > j) &= P(W^0_{m(j)} \neq \emptyset, N > j-1, W^{(z_j,m(j))}_{1,m(j+1)} = \emptyset) \\ &\leq (1-\rho_\infty) P(W^0_{m(j)} \neq \emptyset, N > j-1). \end{aligned}$$

The above is at most $(1-\rho_\infty)P(W^0_{m(j-1)} \neq \emptyset, N > j-1)$, so iterating this, we get

$$P(W^0_{m(j)} \neq \emptyset, N > j) \leq (1-\rho_\infty)^j, \tag{5.8}$$

and if $n > J_0 n_1$, then

$$P(W^0_{2n} \neq \emptyset, N > J_0) \leq P(W^0_{m(J_0)} \neq \emptyset, N > J_0) \leq (1-\rho_\infty)^{J_0} < \delta. \tag{5.9}$$

We need a final preparatory inequality. Using (5.6) and the Markov property, for $n > n_1$ we have

$$\begin{aligned} &P(W^0_{1,2n} \neq \emptyset, W^0_{2n} \cap A = \emptyset) \\ &\quad \leq \delta + \sum_{B \subset A', |B| \geq K_0} P(W^0_{1,2n_1} \cap \pi_1 A' = \pi_1 B) P(x \notin W^{(\pi_1 x, 2n_1)}_{2,2n} \ \forall x \in B) \\ &\quad \leq \delta + \sum_{B \subset A', |B| \geq K_0} P(W^0_{1,2n_1} \cap \pi_1 A' = \pi_1 B) \prod_{x \in B} P(x \notin W^{(\pi_1 x, 2n_1)}_{2,2n}), \end{aligned}$$

the last step by independence of the slab processes. Thus, employing (5.3),

$$\limsup_{n\to\infty} P(W^0_{1,2n} \neq \emptyset, W^0_{2n} \cap A = \emptyset) \leq \delta + (1-\rho_1)^{K_0} < 2\delta. \tag{5.10}$$

We are ready for the final steps. For each $j \leq J_0$ and $n \geq J_0 n_1$, by the Markov property and (5.7),

$$\begin{aligned} &P(W^0_{2n} \neq \emptyset, W^0_{2n} \cap A = \emptyset, N = j) \\ &\quad \leq P(W^0_{m(j)} \neq \emptyset, N > j-1, W^{(z_j,m(j))}_{1,m(j+1)} \neq \emptyset, W^{(z_j,m(j))}_{2n} \cap A = \emptyset) \\ &\quad = \sum_z P(W^0_{m(j)} \neq \emptyset, N > j-1, z_j = z) \times \\ &\qquad P(W^{(z,m(j))}_{1,m(j+1)} \neq \emptyset, W^{(z,m(j))}_{2n} \cap A = \emptyset) \\ &\quad \leq \sum_z P(W^0_{m(j)} \neq \emptyset, N > j-1, z_j = z) \\ &\qquad \times (\delta + P(W^{(z,m(j))}_{1,2n} \neq \emptyset, W^{(z,m(j))}_{2n} \cap A = \emptyset)). \end{aligned}$$

Applying (5.10) and then (5.8) we obtain

$$\limsup_{n\to\infty} P(W^0_{2n} \neq \emptyset, W^0_{2n} \cap A = \emptyset, N = j) \leq 3\delta P(W^0_{m(j)} \neq \emptyset, N > j-1)$$
$$\leq 3\delta(1-\rho_\infty)^{j-1}.$$

It follows that

$$\limsup_{n\to\infty} P(W^0_{2n} \neq \emptyset, W^0_{2n} \cap A = \emptyset)$$
$$\leq \limsup_{n\to\infty} P(W^0_{2n} \neq \emptyset, N > J_0) + 3\delta \sum_{j=1}^{J_0} (1-\rho_\infty)^{j-1}$$
$$\leq \delta + 3\delta/\rho_\infty$$

by using (5.9) and summing the series. This completes the proof. □

Now we follow [11] and Section 6 of [5] in describing a setup which connects our spin-flip systems with the percolation process defined above. Let K, L, T be finite positive constants with $K, L \in \mathbb{N}$, let $r = \frac{1}{16d}$, $Q_\varepsilon = [0, \lceil \varepsilon^{r-1} \rceil]^d \cap \mathbb{Z}^d$, and $Q(L) = [-L, L]^d$. We define a set H of configurations in $\{0,1\}^{\mathbb{Z}^d}$ to be an unscaled version of the set of configurations in $\{0,1\}^{\varepsilon\mathbb{Z}^d}$ of the same name in Section 6 of [5], that is,

$$H = \{\xi \in \{0,1\}^{\mathbb{Z}^d} : |Q_\varepsilon|^{-1} \sum_{y \in Q_\varepsilon} \xi(x+y) \in I^* \text{ for all } x \in Q(L) \cap ([0, \lceil \varepsilon^{r-1} \rceil]^d \cap \mathbb{Z}^d)\}.$$

Here I^* is a particular closed subinterval of $(0,1)$ (it is I^*_η in the notation of Section 6 in [5]). The key property we will need of H is

(5.11) for each $\xi \in H$ there are $y_0, y_1 \in Q(L) \cap \mathbb{Z}^d$ s.t. $\xi(y_i) = i$ for $i = 0, 1$.

This is immediate from the definition and the fact that I^* is a closed subinterval of $(0,1)$. For $z \in \mathbb{Z}^d$, let $\sigma_z : \{0,1\}^{\mathbb{Z}^d} \to \{0,1\}^{\mathbb{Z}^d}$ be the translation map, $\sigma_z(\xi)(x) = \xi(x+z)$ and let $0 < \gamma' < 1$. Recall from Section 3 that for $R \subset \mathbb{R}^d$, $\mathcal{G}(R \times [0,T])$ is the σ-field generated by the points of the graphical construction in the space-time region $R \times [0,T]$. For each $\xi \in H$, G_ξ will denote an event such that

(i) G_ξ is $\mathcal{G}([-KL, KL]^d \times [0,T])$-measurable,
(ii) If $\xi_0 = \xi \in H$ then on G_ξ, $\xi_T \in \sigma_{Le} H$ for all $e \in \{e_1, -e_1, \ldots, e_d, -e_d\}$,
(iii) $P(G_\xi) \geq 1 - \gamma'$ for all $\xi \in H$.

We are now in a position to quote the facts we need from Section 6 of [5], which depend heavily on our assumption $f'(0) > 0$ (and by symmetry, $f'(1) < 0$). This allows us to use Proposition 1.6 of [5] to show that Assumption 1 of that reference is in force and so by a minor modification of Lemma 6.3 of [5] we have the following.

Lemma 5.2. *For any $\gamma' \in (0,1)$ there exists $\varepsilon_1 > 0$ and finite $K \in \mathbb{N}$ such that for all $0 < \varepsilon < \varepsilon_1$ there exist $L, T, \{G_\xi, \xi \in H\}$, all depending on ε, satisfying the basic setup given above.*

Lemma 6.3 of [5] deals with a rescaled process on the scaled lattice $\varepsilon\mathbb{Z}^d$ but here we have absorbed the scaling parameters into our constants T and L and then shifted L slightly so that it is a natural number. In fact L will be of the form $\lceil c\varepsilon^{-1}\log(1/\varepsilon)\rceil$.

Given $\xi = \xi_0 \in \{0,1\}^{\mathbb{Z}^d}$ we define

$$V_n = \{x : (x,n) \in \mathcal{L} \text{ and } \sigma_{-Lx}\xi_{nT} \in H\}. \tag{5.12}$$

Note that $V_n = \emptyset$ and $V_{n+1} \neq \emptyset$ is possible. Theorem A.4 of [11] and its proof imply that there are $\{0,1\}$-valued random variables $\{\theta'(z,n) : (z,n) \in \mathcal{L}\}$ so that if $\{W'_m{}^{(x,n)} : m \in \mathbb{Z}_+, (x,n) \in \mathcal{L}\}$ and $\{\mathcal{C}'(z,n) : (z,n) \in \mathcal{L}\}$ are constructed from $\{\theta'(z,n)\}$ as above, then

$$\text{if } x \in V_n, \text{ then } W'_m{}^{(x,n)} \subset V_m \text{ for all } m \geq n, \tag{5.13}$$

and $\{W'_n\}$ is a $2K$-dependent oriented percolation process, that is,

$$P(\theta'(z_k,n_k) = 1 \mid \theta'(z_j,n_j), j < k) \geq 1 - \gamma' \tag{5.14}$$

whenever $(z_j,n_j), 1 \leq j \leq k$ satisfy $n_j < n_k$, or $n_j = n_k$ and $|z'_j - z'_k| > 2K$, for all $j < k$. The Markov property of ξ^ε allows us to only require $n_j < n_k$ as opposed to $n_k - n_j > 2K$ in the above, as in Section 6 of [5].

Let $\Delta = (2K+1)^{d+1}$. By Theorem B26 of [25], modified as in Lemma 5.1 of [5], if γ' (in Lemma 5.1) is taken small enough so that $1-\gamma = (1-(\gamma')^{1/\Delta})^2 \geq 1/4$, then the $\theta'(z,n)$ can be coupled with iid Bernoulli variables $\theta(z,n)$ such that

$$\begin{aligned} &\theta(z,n) \leq \theta'(z,n) \text{ for all } (z,n) \in \mathcal{L}, \text{ and} \\ &P(\theta(z,n) = 1) = 1 - \gamma. \end{aligned} \tag{5.15}$$

(The simpler condition on γ and γ' in Theorem B26 of [25] and above in fact follows from that in [26] and Lemma 5.1 of [5] by some arithmetic, and the explicit value of Δ comes from the fact that we are now working on $\mathbb{Z}^d$.) If the coupling part of (5.15) holds, then $W_n \subset V_n$ for all n and (5.13) implies

$$x \in V_n \text{ implies } W_m^{(x,n)} \subset V_m \text{ for all } m \geq n. \tag{5.16}$$

Now choose γ' small enough in Lemma 5.2 so that

$$1 - \gamma = (1-(\gamma')^{1/\Delta})^2 > 1 - 6^{-4}. \tag{5.17}$$

We can now verify condition (4.1).

Lemma 5.3. *If ξ^ε is a voter model perturbation satisfying the hypotheses of Theorem 1.2, then there exists $\varepsilon_1 > 0$ and $x_0 \in \mathbb{Z}^d$ such that (4.1) holds for ξ^ε if $\varepsilon < \varepsilon_1$.*

Proof. For γ' as above, let ε_1 be as in Lemma 5.2, so that for $0 < \varepsilon < \varepsilon_1$ all the conclusions of that lemma hold, as well as the setup (5.11)-(5.16), with $\rho_\infty > 0$. There are two main steps in the proof. In the first, we show that if $A \subset 2\mathbb{Z}^d$ is large then for all large n, $\xi^\varepsilon_{2nT} \neq \emptyset$ will imply $V_{2n} \cap A$ is also large (see (5.28) below). To do this, we argue that there is a uniform positive lower bound on $P_\xi(\exists\ z \in V_2$ with $W_m^{(z,2)} \neq \emptyset\ \forall\ m \geq 2)$, $\xi \notin \{\mathbf{0}, \mathbf{1}\}$. Iteration leads to (5.28). In the second step, we consider $A \subset \mathbb{Z}^d$ large, and for $a \in A$ choose points $\ell(a) \in \mathbb{Z}^d$ such that $a \in 2L\ell(a) + Q(L)$. If A is sufficiently large there will be many points $a_i \in A$ which are widely separated. By the first step, for large n, there will be many points $2\ell(a_i) \in V_{2n}$, and for each of these there will be points $y_i^0, y_i^1 \in 2L\ell(a_i) + Q(L)$ such that $\xi^\varepsilon_{2nT}(y_i^0) = 0$ and $\xi^\varepsilon_{2nT}(y_i^1) = 1$. Given these points, it will follow from Lemma 3.4 that there is a uniform positive lower bound on the probabilities of independent events on which $\xi^\varepsilon_t(a_i) = 1$ and $\xi^\varepsilon_t(a_i + x_0) = 0$ for all $t \in [(2n+1)T, (2n+3)T]$. Many of these events will occur, forcing $A(\xi_t, x_0)$ to be large (see (5.30) below). Condition (4.1) now follows easily.

It is convenient to start with two estimates which depend only on the process ξ^ε (and not on the percolation construction). We claim that by Lemma 3.4 with $t = 2T$,

$$\min_{x \in Q(L), k=1,\ldots,d} \inf_{\substack{\xi \in \{0,1\}^{\mathbb{Z}^d}: \\ \xi(x)=1, \xi(x+e_k)=0}} P_\xi(\xi^\varepsilon_{2T} \in H) > 0. \tag{5.18}$$

To see this, note that for small ε, $\xi \in H$ depends only on the coordinates $\xi(x)$, $x \in Q(L+1)$. This means there are disjoint sets $B_0, B_1 \subset Q(L+1)$ so that $\xi^\varepsilon_{2T}(x) = i$ for all $x \in B_i$, $i = 0, 1$ implies $\xi^\varepsilon_{2T} \in H$. If $G(2T, y_0, y_1, B_0, B_1)$ and $\Lambda(y_0, y_1, B_0, B_1)$ are as in Lemma 3.4 with $(y_0, y_1) = (x + e_k, x)$, then for $x \in Q(L)$ the above infimum is bounded below by $P(G(2T, x + e_k, x, B_0, B_1)) > 0$. If $\xi \notin \{\mathbf{0}, \mathbf{1}\}$ there must exist $k \in \{1, \ldots, d\}$ and $x, z \in \mathbb{Z}^d$ with $x \in 2Lz + Q(L)$ and $\xi(x) = 1$, $\xi(x + e_k) = 0$. It now follows from translation invariance that

$$\rho_1 = \inf_{\xi \notin \{\mathbf{0}, \mathbf{1}\}} P_\xi(\exists z \in \mathbb{Z}^d \text{ such that } \xi^\varepsilon_{2T} \circ \sigma_{2Lz} \in H) > 0. \tag{5.19}$$

Let $\rho_2 = \rho_1 \rho_\infty > 0$.

Next, suppose $y_0, y_1, y \in Q(L)$, $B_1 = \{y\}$, $B_0 = \{y + x_0\}$ and $G(T, y_0, y_1, B_0, B_1)$ be as in Lemma 3.4. To also require that ξ^ε_u be constant at $y, y + x_0$ for $u \in [T, 3T]$ we let $\tilde{G}(T, y_0, y_1, B_0, B_1)$ be the event

$$G(T, y_0, y_1, B_0, B_1) \cap \{T_m^z, T_m^{*,z} \notin [T, 3T] \text{ for } z = y, y + x_0 \text{ and all } m \geq 1\}.$$

Note that each $\tilde{G}$ is an intersection of two independent events each with positive probability, and so $P(\tilde{G}) > 0$. Making use of the notation of Lemma 3.4, choose

$\tilde{M} < \infty$ such that

$$\Lambda = \bigcup_{y_0, y_1, y \in Q(L)} \Lambda(y_0, y_1, B_0, B_1) \subset [-\tilde{M}, \tilde{M}]^d$$

and put

$$\tilde{\delta} = \min_{y_0, y_1, y \in Q(L)} P(\tilde{G}(T, y_0, y_1, B_0, B_1)) > 0. \tag{5.20}$$

If $\xi_0^\varepsilon(y_i) = i, i = 0, 1$, then

(5.21)

$$\tilde{G}(T, y_0, y_1, B_0, B_1) \text{ implies } \xi_t^\varepsilon(y) = 1, \xi_t^\varepsilon(y + x_0) = 0 \text{ for all } t \in [T, 3T].$$

We now start the proof of

$$\lim_{K \to \infty} \sup_{\substack{A \subset 2\mathbb{Z}^d \\ |A| \geq K}} \lim_{n \to \infty} P_\xi(\xi_{2nT}^\varepsilon \neq \emptyset \text{ and } V_{2n} \cap A = \emptyset) = 0. \tag{5.22}$$

Fix $\delta > 0$. By (5.2) there exists $K_1 = K_1(\delta) < \infty$ such that if $A \subset 2\mathbb{Z}^d$ with $|A| \geq K_1$ then there exists $n_1 = n_1(A) < \infty$ such that

$$P(W_{2n}^0 \neq \emptyset, W_{2n}^0 \cap A = \emptyset) < \delta \text{ for all } n \geq n_1. \tag{5.23}$$

We may increase n_1 if necessary so that $P(|\mathcal{C}(0,0)| < \infty, W_{2n_1}^0 \neq \emptyset) < \delta$, which implies that

$$P(W_{2n_1}^0 \neq \emptyset, W_{2n}^0 = \emptyset) < \delta \text{ for all } n \geq n_1. \tag{5.24}$$

For $j = 1, 2 \ldots$, let $m(j) = (j-1)(2n_1 + 2)$, and define a random sequence of sites z_j, as follows. If $V_{m(j)+2} = \emptyset$ put $z_j = 0$. If not, choose $z \in V_{m(j)+2}$ with minimal norm (with some convention for ties) and put $z_j = z$. By the Markov property, and (5.19),

$$\inf_{\xi \notin \{\mathbf{0}, \mathbf{1}\}} P_\xi(z_1 \in V_2, |\mathcal{C}(z_1, 2)| = \infty) \geq \rho_2. \tag{5.25}$$

Let

$$N = \inf\{j : z_j \in V_{m(j)+2} \text{ and } W_{m(j+1)}^{z_j, m(j)+2} \neq \emptyset\},$$

and $\mathcal{F}_n$ be the σ-algebra generated by $\mathcal{G}(\mathbb{R}^d \times [0, nT])$ and the $\theta(z, k)$ for $z \in \mathbb{Z}^d, k < n$. It follows from our construction and (5.25) that almost surely on the event $\{\xi_{m(j)T}^\varepsilon \neq \emptyset\}$,

$$\begin{aligned} P_\xi\Big(z_j \in V_{m(j)+2} \text{ and } W_{m(j+1)}^{(z_j, m(j)+2)} = \emptyset \mid \mathcal{F}_{m(j)}\Big) \\ &= P_{\xi_{m(j)T}^\varepsilon}(z_1 \in V_2 \text{ and } W_{2n_1+2}^{(z_1, 2)} = \emptyset) \\ &\leq P_{\xi_{m(j)T}^\varepsilon}(z_1 \in V_2, |\mathcal{C}(z_1, 2)| < \infty) \\ &\leq 1 - \rho_2. \end{aligned}$$

In the last line note that by (5.11) if the initial state is $\mathbf{1}$, the probability is zero as $\mathbf{1}$ is a trap. Since the event on the LHS is $\mathcal{F}_{m(j+1)}$-measurable we may iterate this inequality to obtain

$$P_\xi(\xi^\varepsilon_{m(j)T} \neq \emptyset, N > j) \leq (1-\rho_2)^j. \tag{5.26}$$

Taking $J_0 > 2$ large enough so that $(1-\rho_2)^{J_0} < \delta$, and then $2n > m(J_0+1)$,

$$P_\xi(\xi^\varepsilon_{2nT} \neq \emptyset, V_{2n} \cap A = \emptyset) \leq \delta + \sum_{j=1}^{J_0} P_\xi(\xi^\varepsilon_{m(j)T} \neq \emptyset, V_{2n} \cap A = \emptyset, N = j). \tag{5.27}$$

For $j \leq J_0$, almost surely on the event $\{\xi^\varepsilon_{m(j)T} \neq \emptyset, N > j-1\}$,

$$\begin{aligned}
P_\xi\Big(z_j \in V_{m(j)+2}, &W^{(z_j,m(j)+2)}_{m(j+1)} \neq \emptyset, V_{2n} \cap A = \emptyset \mid \mathcal{F}_{m(j)}\Big) \\
&= P_{\xi^\varepsilon_{m(j)}}\Big(z_1 \in V_2, W^{(z_1,2)}_{2n_1} \neq \emptyset, V_{2n-2n_1} \cap A = \emptyset\Big) \\
&\leq P_{\xi^\varepsilon_{m(j)}}\Big(z_1 \in V_2, W^{(z_1,2)}_{2n_1} \neq \emptyset, W^{(z_1,2)}_{2n-2n_1} \cap A = \emptyset\Big) \\
&\leq \delta + P_{\xi^\varepsilon_{m(j)}}\Big(z_1 \in V_2, W^{(z_1,2)}_{2n-2n_1} \neq \emptyset, W^{(z_1,2)}_{2n-2n_1} \cap A = \emptyset\Big) \\
&\leq 2\delta,
\end{aligned}$$

where the last three inequalities follow from (5.16), (5.24), (5.23), and the fact that $n \geq 2n_1$ by our choice of n above. Combining this bound with (5.27), and then using (5.26), we obtain

$$\begin{aligned}
P_\xi(\xi^\varepsilon_{2nT} \neq \emptyset, V_{2n} \cap A = \emptyset) &\leq \delta + 2\delta \sum_{j=1}^{J_0} P_\xi(\xi^\varepsilon_{m(j)T} \neq \emptyset, N > j-1) \\
&\leq \delta + 2\delta \sum_{j=1}^{J_0} (1-\rho_2)^{j-1} \\
&\leq \delta + 2\delta/\rho_2.
\end{aligned}$$

This establishes (5.22), which along with the argument proving (4.4) implies that for any $K_0 < \infty$,

$$\lim_{K\to\infty} \sup_{\substack{A \subset 2\mathbb{Z}^d \\ |A| \geq K}} \limsup_{n\to\infty} P_\xi(\xi^\varepsilon_{2nT} \neq \emptyset, |V_{2n} \cap A| \leq K_0) = 0. \tag{5.28}$$

Now fix $K_0 < \infty$ so that $(1-\tilde{\delta})^{K_0} < \delta$. By (5.28) there exists $K_1 < \infty$ such that for $A' \subset 2\mathbb{Z}^d$ satisfying $|A'| \geq K_1$, there exists $n_1(A')$ so that

$$P_\xi(\xi^\varepsilon_{2nT} \neq \emptyset \text{ and } |V_{2n} \cap A'| \leq K_0) < \delta \quad \text{if } n \geq n_1(A'). \tag{5.29}$$

For $a \in \mathbb{Z}^d$ let $\ell(a)$ be the minimal point in some ordering of $\mathbb{Z}^d$ such that $a \in 2L\ell(a) + Q(L)$. For $A \subset \mathbb{Z}^d$ let $\ell(A) = \{\ell(a), a \in A\}$. With K_0, K_1 as

above, choose $K_2 < \infty$ so that if $A \subset \mathbb{Z}^d$ and $|A| \geq K_2$ then $\ell(A)$ contains K_1 points, $\ell(a_1), \ldots, \ell(a_{K_1})$, such that $|\ell(a_i) - \ell(a_j)|2L \geq 4\tilde{M}$ for $i \neq j$. The regions $2L\ell(a_1)+[-\tilde{M}, \tilde{M}]^d, \ldots, 2L\ell(a_{K_1})+[-\tilde{M}, \tilde{M}]^d$ are pairwise disjoint. Let $A' = \{2\ell(a_1), \ldots, 2\ell(a_{K_1})\} \subset 2\mathbb{Z}^d$.

Now suppose $t \in [(2n+1)T, (2n+3)T]$ for some integer $n \geq n_1(A')$. By (5.29), on the event $\{|\xi^\varepsilon_{2nT}| > 0\}$, except for a set of probability at most δ, V_{2n} will contain at least K_0 points of A'. If $2\ell(a_i)$ is such a point, then by the definitions of V_{2n} and H, there will exist points $y^i_0, y^i_1 \in 2L\ell(a_i) + Q(L)$ such that $\xi^\varepsilon_{2nT}(y^i_0) = 0, \xi^\varepsilon_{2nT}(y^i_1) = 1$. Conditional on this, by (5.20) and (5.21), the probability that $\xi^\varepsilon_t(a_i) = 1, \xi^\varepsilon_t(a_i + x_0) = 0$ is at least $\tilde{\delta}$. By independence of the Poisson point process on disjoint space-time regions, it follows that

$$P(\xi^\varepsilon_{2nT} \neq \emptyset \text{ and } A(x_0, \xi^\varepsilon_t) = \emptyset) < \delta + (1 - \tilde{\delta})^{K_0}, \tag{5.30}$$

and therefore since $t > 2nT$,

$$P(\xi^\varepsilon_t \neq \emptyset \text{ and } A(x_0, \xi^\varepsilon_t) = \emptyset) < \delta + (1 - \tilde{\delta})^{K_0} < 2\delta,$$

the last by our choice of K_0. This proves (4.1).

Finally, (5.25) implies by (5.16), (5.11), the definition of V_n, and the fact that $\mathbf{1}$ is a trap by Lemma 2.1, that

$$\inf_{\xi \neq \mathbf{0}} P_\xi(\xi^\varepsilon_t \neq \emptyset \ \forall\ t \geq 0) \geq \rho_2. \tag{5.31}$$

This will be used below. □

Proof of Theorem 1.2. We verify the assumptions of Proposition 4.1. It follows from Lemma 2.1 and (1.15), $c_\varepsilon(x, \xi)$ is symmetric and ζ^ε_t, the annihilating dual of ξ^ε_t, is parity preserving. By Corollary 3.3 (which applies by Remark 5) there exists $\varepsilon_3 > 0$ such that if $0 < \varepsilon < \varepsilon_3$ then ζ^ε_t is irreducible. By Lemmas 4.2 and 5.3 (and the proof of the latter), there exists $0 < \varepsilon_4 < \varepsilon_3$ such that if $0 < \varepsilon < \varepsilon_4$ then (4.1), (4.2) and (5.31) hold for ξ^ε_t.

Assume now that $0 < \varepsilon < \varepsilon_4$. It remains to check that the dual growth condition (2.7) (the conclusion of Lemma 2.2) holds, and to do this it suffices by Remark 4 to show that (2.8) for ξ^ε holds. By (4.1) and (5.31) there is a $\delta_1 > 0$, $t_0 < \infty$ and $A \in Y$ so that for all $t \geq t_0$ (with $\xi^\varepsilon_0 = 1_{\{0\}}$),

$$P(\xi^\varepsilon_t(a) = 1 \text{ for some } a \in A) \geq P(A(x_0, \xi^\varepsilon_t) \neq \emptyset) \geq \delta_1.$$

Next apply (3.9), translation invariance and the Markov property to conclude that for t as above

$$\begin{aligned} P(\xi^\varepsilon_{t+1}(0) = 1) &\geq E(1(\xi^\varepsilon_t(a) = 1 \text{ for some } a \in A) P_{\xi^\varepsilon_t}(\xi^\varepsilon_1(0) = 1)) \\ &\geq \delta_1 \min_{a \in A} \inf_{\xi^\varepsilon_0 : \xi^\varepsilon_0(0) = 1} P_{\xi^\varepsilon_0}(\xi^\varepsilon_1(-a) = 1) \geq \delta_2 > 0. \end{aligned}$$

This proves (2.8), and now all the assumptions of Proposition 4.1 have now been verified for ξ^ε_t if $0 < \varepsilon < \varepsilon_4$, and thus the weak limit (4.3) also holds. Finally, by (1.8) this result implies the full complete convergence theorem with coexistence if $|\hat{\xi}^\varepsilon_0| = \infty$. If $|\hat{\xi}^\varepsilon_0| < \infty$, then $|\xi^\varepsilon_0| = \infty$ and the result now follows by the symmetry of ξ^ε (recall Lemma 2.1). □

6. Proof of Theorem 1.1

Let us check that $LV(\alpha)$, $\alpha \in (0,1)$, is cancellative. (This was done in [27] for the case $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$ for $\mathcal{N}$ satisfying (1.24).) For the more general setting here, we assume $p(x)$ satisfies (1.1), and allow any $d \geq 1$. We first observe that if $c(x,\xi)$ has the form given in (1.16), then it follows from (1.17) that

$$c(0,\xi) = k_0 \sum_{A \in Y} q_0(A)\ \tfrac{1}{2}[1 - (2\xi(0) - 1)H(\xi, A)].$$

From this it is clear that the sum of two positive multiples of cancellative rate functions is cancellative. It follows from a bit of arithmetic that if (1.5) holds, then $LV(\alpha)$ with $\varepsilon^2 = 1 - \alpha > 0$ has flip rates

$$c_{\mathrm{LV}}(x,\xi) = \alpha c_{\mathrm{VM}}(x,\xi) + \varepsilon^2 f_0 f_1(x,\xi).$$

We have already noted that c_{VM} is cancellative, and so by the above we need only check that $c^*(x,\xi) = f_0(x,\xi)f_1(x,\xi)$ is cancellative.

To do this we let $p^{(2)}(0) = \sum_{x \in \mathbb{Z}^d} (p(x))^2$, $k_0 = (1 - p^{(2)}(0))/2$, $q_0(A) = 0$ if $|A| \neq 3$, and

$$q_0(\{0,x,y\}) = k_0^{-1} p(x)p(y) \text{ if } 0,x,y \text{ are distinct.}$$

Note that $\sum_{A \in Y} q_0(A) = 1$ because (recall that $p(0) = 0$)

$$\sum_{\{x,y\}} q_0(\{0,x,y\}) = \frac{1}{2k_0} \sum_{x \neq y} p(x)p(y) = \frac{1}{2k_0}(1 - p^{(2)}(0)) = 1.$$

Also, for $0,x,y$ distinct,

$$\tfrac{1}{2}[1 - (2\xi(0) - 1)H(\xi, \{0,x,y\})] = 1\{\xi(x) \neq \xi(y)\}.$$

With these facts it is easy to see that

$$k_0 \sum_{A \in Y} q_0(A)\ \tfrac{1}{2}[1 - (2\xi(0) - 1)H(\xi, A)]$$
$$= \sum_{x,y} p(x)p(y) 1\{\xi(x) \neq \xi(y)\} = f_0(0,\xi)f_1(0,\xi),$$

proving $c^*(x,\xi) = f_0(x,\xi)f_1(x,\xi)$ is cancellative and hence so is $LV(\alpha)$.

Although we won't need it, we calculate the parameters of the branching annihilating dual. Adding in the voter model, we see that they are

$$k_0 = \alpha + (1-\alpha)\frac{1 - p^2(0)}{2},\ q_0(\{y\}) = \frac{\alpha}{k_0} p(y),\ q_0(\{0,y,z\}) = \frac{1-\alpha}{k_0} p(y)p(z),$$

and $q_0(A) = 0$ otherwise. One can see from this that ζ_t, the dual of $LV(\alpha)$, describes a system of particles evolving according to the following rules: (i) a

particle at x jumps to y at rate $\alpha p(y-x)$, (ii) a particle at x creates two particles and sends them to y, z at rate $(1-\alpha)p(y-x)p(z-x)$, (ii) if a particle attempts to land on another particle then the two particles annihilate each other.

Assume $d \geq 3$. The function $f(u)$ as shown in Section 1.3 of [5] is a cubic, and under the assumption (1.5) reduces to $f(u) = 2p_3(1-\alpha)u(1-u)(1-2u)$, where p_3 is a certain (positive) coalescing random walk probability. Thus $f'(0) > 0$, so the complete convergence theorem with coexistence for $LV(\alpha)$ for α sufficiently close to one follows from Theorem 1.2.

Now suppose $d = 2$. It suffices to prove an analogue of Lemma 5.2 as the above results will then allow us to apply the proof of Theorem 1.2 in the previous section to give the result. As the results of [5] do not apply we will use results from [6] instead and proceed as in Section 4 of [8]. Instead of (1.2), we only require (as was the case in [6])

$$\sum_{x\in\mathbb{Z}^2} |x|^3 p(x) < \infty. \tag{6.1}$$

We will need some notation from [6]. For $N > 1$ let $\xi^{(N)}$ be the $LV(\alpha_N)$ process where

$$\alpha_N = 1 - \frac{(\log N)^3}{N},$$

and consider the rescaled process, $\xi_t^N(x) = \xi_{Nt}^{(N)}(x\sqrt{N})$, for $x \in S_N = \mathbb{Z}^2/\sqrt{N}$. The associated process taking values in $M_F(\mathbb{R}^2)$ (the space of finite measures on the plane with the weak topology) is

$$X_t^N = \frac{\log N}{N} \sum_{x\in S_N} \xi_t^N(x)\delta_x. \tag{6.2}$$

For parameters $K, L' \in \mathbb{N}$, $K > 2$ and $L' > 3$, which will be chosen below, we let $\underline{\xi}_t^N(x) \leq \xi_t^N(x)$, $x \in S_N$ be a coupled particle system where particles are "killed" when they exit $(-KL', KL')^2$, as described in Proposition 2.1 of [8]. (Here a particle corresponds to a 1.) In particular $\underline{\xi}_t^N(x) = 0$ for all $|x| \geq KL'$. $\underline{X}_t^N$ is defined as in (6.2) with $\underline{\xi}^N$ in place of ξ^N.

We will need to keep track of some of the dependencies in the constant $C_{8.1}$ in Lemma 8.1 of [6]. As in that result, B^N is a rate $N\alpha_N = N - (\log N)^3$ random walk on S_N with step distribution $p_N(x) = p(x\sqrt{N})$, $x \in S_N$, starting at the origin.

Lemma 6.1. *There are positive constants c_0 and δ_0, and a non-decreasing function $C_0(\cdot)$ so that if $t > 0$, $K, L' \in \mathbb{N}$, $K > 2$ and $L' > 3$, and $X_0^N = \underline{X}_0^N$ is supported on $[-L', L']^2$, then*

$$E(X_t^N(1) - \underline{X}_t^N(1)) \leq X_0^N(1)\Big[c_0 e^{c_0 t} P(\sup_{s\leq t}(|B_s^N| > (K-1)L' - 3) + C_0(t)(1 \vee X_0^N(1))(\log N)^{-\delta_0}\Big]. \tag{6.3}$$

Proof. This is a simple matter of keeping track of the t-dependency in some of the constants arising in the proof of Lemma 8.1 in [6]. □

Recall from Theorem 1.5 of [6] that if $X_0^N \to X_0$ in $M_F(\mathbb{R}^2)$, then $\{X^N\}$ converges weakly in $D(\mathbb{R}_+, M_F(\mathbb{R}^2))$ to a two-dimensional super-Brownian motion, X, with branching rate $4\pi\sigma^2$, diffusion coefficient σ^2 and drift $\eta > 0$ (write X is $SBM(4\pi\sigma^2, \sigma^2, \eta)$), where η is the constant K in (6) of [6] (not to be confused with our parameter K). See (MP) in Section 1 of [6] for a precise definition of SBM. The important point for us is that the positivity of η will mean that the supercritical X will survive with positive probability, and on this set will grow exponentially fast.

We next prove a version of Proposition 4.2 of [8] which when symmetrized is essentially a scaled version of the required Lemma 5.2. To be able to choose γ' as in (5.17), so that we may apply Lemma 5.1 of [5], we will have to be more careful with the selection of constants in the proof of Proposition 4.2 in the above reference. We start by choosing $c_1 > 0$ so that

$$(1 - e^{-c_1})^2 > 1 - 6^{-4}, \tag{6.4}$$

and then setting

$$\gamma'_K = e^{-c_1(2K+1)^3}.$$

Lemma 6.2. *There are $T' > 1$, $L', K, J' \in \mathbb{N}$ with $K > 2$, $L' > 3$, and $\varepsilon_1 \in (0, \frac{1}{2})$ such that if $0 < 1 - \alpha < \varepsilon_1$, $N > 1$ is chosen so that $\alpha = 1 - \frac{(\log N)^3}{N}$, and $I_{\pm e_i} = \pm 2L' e_i + [-L', L']^2$, then*

$$\underline{X}_0^N([-L', L']^2) \geq J' \text{ implies } P(\underline{X}_{T'}^N(I_{\pm e_i}) \geq J' \text{ for } i = 1, 2) \geq 1 - \gamma'_K.$$

Proof. By the monotonicity of $\underline{X}^N$ in its initial condition (Proposition 2.1(b) of [8] and the monotonicity of $LV(\alpha)$ discussed, for example, in Section 1 of [8]) we may assume that $\underline{X}_0^N(\mathbb{R}^2 \setminus [-L', L']^2) = 0$ and $\underline{X}_0^N([-L', L']^2) \in [J', 2J']$, where L' and J' are chosen below.

We will choose a number of constants which depend on an integer $K > 2$ and will then choose K large enough near the end of the proof. Assume $B = (B^1, B^2)$ is a 2-dimensional Brownian motion with diffusion parameter σ^2, starting at x under P_x and fix $p > \frac{1}{2}$. Set

$$T' = c_2 K^{2p}, \tag{6.5}$$

where a short calculation shows that if c_2 is chosen large enough, depending on σ^2 and η, then for any $K > 2$,

$$e^{\eta T'/2} \inf_{|x| \leq K^p} P_x(B_1 \in [K^p, 3K^p]^2) \geq 5. \tag{6.6}$$

Now put $L' = K^p \sqrt{T'}$, increasing c_2 slightly so that $L' \in \mathbb{N}$. If $I = [-L', L']^2$ and X is the limiting super-Brownian motion described above, then as in Lemma 12.1(b) of [17], there is a $c_3(K)$ so that

$$\forall J' \in \mathbb{N} \text{ and } i \leq 2, \text{ if } X_0(I) \geq J', \text{ then } P(X_{T'}(I_{\pm e_i}) < 4J') \leq c_3/J'. \tag{6.7}$$

Next choose $J' = J'(K) \in \mathbb{N}$ so that

$$\frac{c_3}{J'} \le \frac{\gamma'_K}{100}.$$

As in Lemma 4.4 of [8], the weak convergence of X^N to X and (6.7) show that for $N \ge N_1(K)$,

$$\forall i \le 2 \text{ if } X_0^N(I) \ge J', \text{ then } P(X_{T'}^N(I_{\pm e_i}) < 4J') \le \frac{\gamma'_K}{50}. \tag{6.8}$$

Next use Lemma 6.1, the fact that $X_{T'}^N - \underline{X}_{T'}^N$ is a non-negative measure, and Donsker's theorem to see that there is a $c_4 > 0$ and an $\varepsilon_N = \varepsilon_N(K) \to 0$ as $N \to \infty$, so that for any $i \le 2$,

$$\begin{aligned}
&P(X_{T'}^N(I_{\pm e_i}) - \underline{X}_{T'}^N(I_{\pm e_i}) \ge 2J') \\
&\quad \le \frac{X_0^N(1)}{2J'}\Big[c_0 e^{c_0 T'}(P_0(\sup_{s\le T'}|B_s| > (K-1)L' - 3) + \varepsilon_N) + C_0(T')(1 \vee X_0^N(1))(\log N)^{-\delta_0}\Big] \\
&\quad \le \Big[c'_0 e^{c_0 T'}(\exp(-c_4 K^{2+2p}) + \varepsilon_N) + C_0(T')2J'(\log N)^{-\delta_0}\Big],
\end{aligned}$$

where the fact that $X_0^N(1) \le 2J'$ and the definition of L' are used in the last line. It follows that for $K \ge K_0$ and $N \ge N_2(K)$, the above is bounded by

$$2c'_0 e^{c_0 T'} \exp(-c_4 K^{2+2p}) \le 2c'_0 e^{c_5 K^{2p} - c_4 K^{2+2p}} \le \frac{\gamma'_K}{50}. \tag{6.9}$$

The fact that $p > \frac{1}{2}$ is used in the last inequality. We finally choose $K \in \mathbb{N}^{>2}$, $K \ge K_0$. Therefore the bounds in (6.8) and (6.9) show that for $N \ge N_1(K) \vee N_2(K)$ and $i \le 2$,

$$\begin{aligned}
P(\underline{X}_{T'}^N(I_{\pm e_i}) < 2J') &\le P(X_{T'}^N(I_{\pm e_i}) \le 4J') + P(X_{T'}^N(I_{\pm e_i}) - \underline{X}_{T'}^N(I_{\pm e_i}) \ge 2J') \\
&\le \frac{\gamma'_K}{25}.
\end{aligned}$$

Sum over the 4 choices of $\pm e_i$ to prove the required result because the condition on N is implied by taking $1 - \alpha = (\log N)^3/N$ small enough. □

Completion of Proof of Theorem 1.1. By symmetry we have an analogue of the above Lemma with 0's in place of 1's. Let α and N be as in Lemma 6.2. Now undo the scaling and set $L = \sqrt{N}L'$, $J = \frac{N}{\log N}J'$ and $T = T'N$. Slightly abusing our earlier notation we let $\underline{\xi}_t \le \xi_t^{(N)}$ be the unscaled coupled particle system where particles are killed upon exiting $(-KL, KL)^2$ and let $\tilde{I}_{\pm e_i} = \pm e_i L + [-L, L]^2$. We define

$$G_\xi = \{\underline{\xi}_T(\tilde{I}_{\pm e_i}) \ge J, \hat{\underline{\xi}}_T(\tilde{I}_{\pm e_i}) \ge J \text{ for } i = 1, 2\},$$

where $\underline{\xi}_0 = \xi$. Lemma 6.2 gives the conclusion of Lemma 5.2 with $\varepsilon = 1 - \alpha$, $\gamma' = \gamma'_K$ and now with

$$H = \{\xi \in \{0,1\}^{\mathbb{Z}^d} : \xi([-L, L]^d) \ge J, \hat{\xi}([-L, L]^d) \ge J\}.$$

Note that by (6.4) and the definition of γ'_K, we have $1-\gamma > 1-6^{-4}$ where γ is as in (5.17). The definition of $\underline{\xi}$ gives the required measurability of G_ξ. Note that H depends only on $\{\xi(x) : x \in [-L,L]^d\}$, and $\xi \in H$ implies $\xi(x)=1$ and $\xi(x')=0$ for some $x,x' \in [-L,L]^d$. These are the only properties of H used in the previous proof. Finally it is easy to adjust the parameters so that $L \in \mathbb{N}$ as in Lemma 5.2. One way to do this is to modify (6.7) so the conclusion of Lemma 6.2 becomes

$$\underline{X}_0^N(I') \geq J' \text{ implies } P(\underline{X}_T^N(I'_{\pm e_i}) \geq J' \ \text{ for } i=1,2) \geq 1-\gamma'_K,$$

where $I' = [-L'-1, L'+1]^2$ and $I'_{\pm e_i} = \pm 2L'e_i + [-L'+1, L'-1]^2$. Then for N large (in addition to the constraints above, $N \geq 9$ will do) one can easily check that the above argument is valid with $L = \lfloor \sqrt{N}L' \rfloor \in \mathbb{N}$. Therefore, with the conclusion of this version of Lemma 5.2 in hand, the result for $d=2$ now follows as in the proof of Theorem 1.2.

Remark 9. *The above argument works equally well for $LV(\alpha)$ for $d \geq 3$ even without assuming (6.1). Only a few constants need to be altered–eg. $p=(d-1)/2$ and $\gamma'_K = e^{c_1(2K+1)^{d+1}}$. More generally the argument is easily adjusted to give the result for the general voter model perturbations in Theorem 1.2 (for $d \geq 3$) without assuming (1.2), provided the particle systems are also attractive. This last condition is needed to use the results in [8].*

7. Proofs of Theorems 1.3 and 1.4

Proof of Theorem 1.3. Let ξ_t be the affine voter model with parameter $\alpha \in (0,1)$, and $d \geq 3$. If $\varepsilon^2 = 1-\alpha$, then the rate function of ξ is of the form in (1.10) and (1.11) where

$$h_i(x,\xi) = -f_i(x,\xi) + 1(\xi(y)=i \text{ for some } y \in \mathcal{N}). \tag{7.1}$$

Taking $Z_1,\ldots,Z_{N_0}$ to be the distinct points in $\mathcal{N}$ we see that $AV(\alpha)$ is a voter model perturbation. The fact that $c_{\text{TV}}(x,\xi)$ is cancellative was established in Section 2 of [3], and so, as for $LV(\alpha)$, we may conclude that $AV(\alpha)$ is a cancellative process. It is easy to check that $c_{\text{AV}}(x,\xi)$ is not a pure voter model rate function, so the only remaining condition of Theorem 1.2 to check is $f'(0) > 0$.

To compute $f(u)$, let $\{B_u^x, u \geq 0, x \in \mathbb{Z}^d\}$ be a system of coalescing random walks with step distribution $p(x)$, and put $A_t^F = \{B_t^x, x \in F\}$, $F \in Y$. The slight abuse of notation $|A_\infty^F| = \lim_{t\to\infty} |A_t^F|$ is convenient. If $\xi_0(x)$ are iid Bernoulli with $E(\xi_0(x)) = u$, and $F_0, F_1 \in Y$ are disjoint, then (see (1.26) in [5])

$$\begin{aligned} &\Big\langle \xi(y)=0 \ \forall\, y \in F_0,\ \xi(x)=1 \ \forall\, x \in F_1 \Big\rangle_u \\ &\qquad = \sum_{i,j} (1-u)^i u^j P(|A_\infty^{F_0}|=i, |A_\infty^{F_1}|=j, |A_\infty^{F_0\cup F_1}|=i+j). \end{aligned} \tag{7.2}$$

From (1.23) and (7.1) we have $f(u) = G_0(u) - G_1(u)$, where

$$G_0(u) = \Big\langle 1\{\xi(0)=0\}(-f_1(0,\xi) + 1\{\xi(y) \neq 0 \text{ for some } y \in \mathcal{N}\})\Big\rangle_u,$$
$$G_1(u) = \Big\langle 1\{\xi(0)=1\}(-f_0(0,\xi) + 1\{\xi(y) \neq 1 \text{ for some } y \in \mathcal{N}\})\Big\rangle_u.$$

If $c_0 = \sum_e p(e)P(|A_\infty^{\{0,e\}}| = 2)$, then the assumption that $0 \notin \mathcal{N}$ and (7.2) imply

$$\begin{aligned} G_0(u) &= -c_0 u(1-u) + \Big\langle 1\{\xi(0)=0\}\Big\rangle_u - \Big\langle 1\{\xi(0)=\xi(y)=0 \text{ for all } y \in \mathcal{N}\}\Big\rangle_u \\ &= -c_0 u(1-u) + 1 - u - \sum_{j=1}^{|\mathcal{N}|+1} (1-u)^j P(|A_\infty^{\mathcal{N}\cup\{0\}}| = j). \end{aligned}$$

Similarly,

$$G_1(u) = -c_0 u(1-u) + u - \sum_{j=1}^{|\mathcal{N}|+1} u^j P(|A_\infty^{\mathcal{N}\cup\{0\}}| = j).$$

Therefore if $A = |A_\infty^{\mathcal{N}\cup\{0\}}|$ we obtain

$$\begin{aligned} f'(0) = G_0'(0) - G_1'(0) &= -1 + \sum_{j=1}^{|\mathcal{N}|+1} jP(A=j) - 1 + P(A=1) \\ &= E(A - 1 - 1(A>1)). \end{aligned}$$

Note that since A is $\mathbb{N}$-valued, we have $A - 1 - 1(A > 1) \geq 0$ with equality holding iff $A \in \{1,2\}$. Hence to show $f'(0) > 0$ it suffices to establish that $P(A > 2) > 0$. But since $|\mathcal{N} \cup \{0\}| \geq 3$ by the symmetry assumption on $\mathcal{N}$, the required inequality is easy to see by the transience of the random walks B_u^x. The complete convergence theorem with coexistence holds if $\varepsilon > 0$ is small enough, depending on $\mathcal{N}$, by Theorem 1.2. □

Proof of Theorem 1.4. Let η_t^θ be the geometric voter model with rate function given in (1.26). Then η_t^θ is cancellative for all $\theta \in [0,1]$ (see Section 2 of [3]), and it is clear that η_t^θ is not a pure voter model for $\theta < 1$. (The latter follows from the fact that $q_0(A) > 0$ for any odd subset of $\mathcal{N} \cup \{0\}$.) The next step is to check that η_t^θ is a voter model perturbation. Clearly $\mathbf{0}$ is a trap. If we set $\varepsilon^2 = 1 - \theta$ and $a_j = c(0,\xi)$ for $\xi(0) = 0$ and $\sum_{x\in\mathcal{N}} \xi(x) = j$, then

$$\begin{aligned} a_j &= \Big[\sum_{k=1}^{j} \binom{j}{k} (-\varepsilon^2)^k\Big] \Big/ \Big[\sum_{k=1}^{|\mathcal{N}|} \binom{|\mathcal{N}|}{k} (-\varepsilon^2)^k\Big] \\ &= \frac{j\varepsilon^2 - \binom{j}{2}\varepsilon^4 + O(\varepsilon^6)}{|\mathcal{N}|\varepsilon^2 - \binom{|\mathcal{N}|}{2}\varepsilon^4 + O(\varepsilon^6)}, \end{aligned}$$

where $\binom{j}{2} = 0$ if $j = 1$. A straightforward calculation (we emphasize that c_{VM} and f_0, f_1 are defined using $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$ which satisfies (1.1) and (1.2)) now shows that

$$c_{\mathrm{GV}}(x,\xi) = c_{\mathrm{VM}}(x,\xi) + \varepsilon^2 \frac{|\mathcal{N}|}{2} f_0(x,\xi) f_1(x,\xi) + O(\varepsilon^4) \text{ as } \varepsilon \to 0, \tag{7.3}$$

where the $O(\varepsilon^4)$ term is uniform in ξ and may be written as a function of $f_1(0,\xi)$. It follows from Proposition 1.1 of [5] and symmetry that ξ^ε is a voter model perturbation.

To apply Theorem 1.2 it only remains to check that $f'(0) > 0$, where

$$\begin{aligned} f(u) &= \Big\langle \Big((1-2\xi(0)\Big) f_1(0,\xi) f_0(0,\xi) \Big\rangle_u \\ &= \sum_{x,y} p(x)p(y) \Big\langle \Big((1-2\xi(0))\xi(x)(1-\xi(y)\Big) \Big\rangle_u . \end{aligned}$$

Using (7.2) it is easy to see that for $x, y, 0$ distinct,

$$\begin{aligned} \Big\langle \xi(x)(1-\xi(y)) \Big\rangle_u &= u(1-u) P(|A_\infty^{\{x,y\}}| = 2), \\ \Big\langle \xi(0)\xi(x)(1-\xi(y)) \Big\rangle_u &= u(1-u) P(|A_\infty^{\{0,x\}}| = 1, |A_\infty^{\{0,x,y\}}| = 2) \\ &\quad + u^2(1-u) P(|A_\infty^{\{0,x,y\}}| = 3) \end{aligned}$$

If we plug the decomposition ($x, y, 0$ still distinct)

$$\begin{aligned} P(|A_\infty^{\{x,y\}}| = 2) &= P(|A_\infty^{\{0,x\}}| = 1, |A_\infty^{\{x,y\}}| = 2) \\ &\quad + P(|A_\infty^{\{0,y\}}| = 1, |A_\infty^{\{x,y\}}| = 2) + P(|A_\infty^{\{0,x,y\}}| = 3) \end{aligned}$$

into the above we find that

$$f(u) = u(1-u)(1-2u) \sum_{x,y} p(x)p(y) P(|A_\infty^{\{0,x,y\}}| = 3), \text{ and}$$

and thus $f'(0) = \sum_{x,y} p(x)p(y) P(|A_\infty^{\{0,x,y\}}| = 3) > 0$ as required. □

Acknowledgement. We thank an anonymous referee for a careful reading of the paper and making a number of helpful suggestions which have improved the presentation.

References

[1] Bramson, M., Ding, W. and Durrett, R. (1991) Annihilating branching processes. *Stochas. Process. Appl.* **37**, 1-17.

[2] Bramson, M. and Neuhauser, C. (1994) Survival of one-dimensional cellular automata under random perturbations. *Ann. Probab.* **22**, 244-263.

[3] Cox, J.T., and Durrett, R. (1991) Nonlinear voter models. *Random Walks, Brownian Motion and Interacting Particle Systems*, R. Durrett and H. Kesten editors, 203-213, Birkhauser.

[4] Cox, J.T., Durrett, R. and Perkins, E.A. (2000) Rescaled voter models converge to super- Brownian motion. *Ann. Probab.* **28** 185-234.

[5] Cox, J.T., Durrett, R. and Perkins, E.A. (2011) Voter model perturbations and reaction diffusion equations. To appear in *Astérisque*.

[6] Cox, J.T., Merle, M. and Perkins, E.A. (2010) Co-existence in a two-dimensional Lotka-Volterra model. *Electron. J. Probab.* **15**, 1190-1266.

[7] Cox, J.T. and Perkins, E.A. (2005) Rescaled Lotka-Volterra models converge to super-Brownian motion. *Ann. Probab.* **33** 904-947.

[8] Cox, J.T. and Perkins, E.A. (2007) Survival and coexistence in stochastic spatial Lotka-Volterra models. *Prob. Theory Rel. Fields* **139** 89-142.

[9] Cox, J.T. and Perkins, E.A. (2008) Renormalization of the Two-dimensional Lotka-Volterra Model. *Ann. Appl. Prob.* **18** 747-812.

[10] Durrett, R., (1984) Oriented percolation in two dimensions. *Ann. Probab.* **12**, 999–1040.

[11] Durrett, R. (1995) Ten Lectures on Particle Systems, in *École d'Été de Probabilités de Saint Flour*, Lecture Notes Math.,**1608** pages 97-201, Springer-Verlag, Berlin.

[12] Durrett, R. (2009) Coexistence in stochastic spatial models. *Ann. Appl. Probab.* **19**, 477–496.

[13] Durrett, R. (2005). *Probability: Theory and Examples, 3rd Ed.*, Thomson, Belmont, CA.

[14] Durrett, R., and Griffeath, D. (1982) Contact processes in several dimensions. *Z. fur Wahr.*, **53**, 535-552.

[15] Durrett, R. and Neuhauser, C. (1991) Epidemics with recovery in $d = 2$. *Ann. Appl. Probab.* **1**, 189–206.

[16] Durrett, R. and Neuhauser, C. (1994) Particle systems and reaction-diffusion equations. *Ann. Probab.* **22**, 289–333.

[17] Durrett, R. and Perkins, E. (1999) Rescaled contact processes converge to super-Brownian motion for $d \geq 2$. *Prob. Theory Rel. Fields.* **114**, 309-399.

[18] Ethier, S.N. and Kurtz, T.G. (1986) *Markov Processes, Characterization and Convergence*. John Wiley and Sons, New York.

[19] Griffeath, D. (1978) Limit theorems for nonergodic set-valued Markov processes. *Ann. Probab.* **6** 379-387.

[20] Griffeath, D. (1979) Additive and cancellative interacting particle systems. *Springer Lecture Notes in Math.* **714**.

[21] Griffeath, D. (1981) The basic contact process. *Stoch. Proc. Appl.* **11**, 151-185.

[22] Handjani, S. (1999) The complete convergence theorem for coexistent voter models. *Ann. Probab.* **27** 226-245.

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

[24] Liggett, T.M. (1994) Coexistence in threshold voter models. *Ann. Probab.* **22** 764-802.

[25] Liggett, T.M. (1999). *Interacting Particle Systems PART 2*, Springer-Verlag, New York.

[26] Liggett, T.M., Schonmann, R.H., Stacey, A.M. (1997) Domination by product measures. *Ann. Probab.* **25** 71-95.

[27] Neuhauser, C. and Pacala, S.W. (1999) An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann. Appl. Probab.* **9** 1226-1259.

[28] Sturm, A. and Swart, J. (2008) Voter models with heterozygosity selection. *Ann. Appl. Probab.* **18** 59-99.