STOCHASTIC SPATIAL MODEL OF PRODUCER–CONSUMER SYSTEMS ON THE LATTICE

N. LANCHIER,* *Arizona State University*

Abstract

The objective of this paper is to give a rigorous analysis of a stochastic spatial model of producer–consumer systems that has been recently introduced by Kang and the author to understand the role of space in ecological communities in which individuals compete for resources. Each point of the square lattice is occupied by an individual which is characterized by one of two possible types, and updates its type in continuous time at rate 1. Each individual being thought of as a producer and consumer of resources, the new type at each update is chosen at random from a certain interaction neighborhood according to probabilities proportional to the ability of the neighbors to consume the resource produced by the individual to be updated. In addition to giving a complete qualitative picture of the phase diagram of the spatial model, our results indicate that the nonspatial deterministic mean-field approximation of the stochastic process fails to describe the behavior of the system in the presence of local interactions. In particular, we prove that, in the parameter region where the nonspatial model displays bistability, there is a dominant type that wins regardless of its initial density in the spatial model, and that the inclusion of space also translates into a significant reduction of the parameter region where both types coexist.

Keywords: Interacting particle system; voter model; Richardson model; threshold contact process

2010 Mathematics Subject Classification: Primary 60K35
Secondary 91A22

1. Introduction

To understand the role of space in ecological communities in which individuals compete for resources, Kang and the author [10] recently introduced a stochastic spatial model of producer–consumer systems on the lattice. Based on their numerical simulations of this spatial model and their analytical results for its nonspatial deterministic mean-field approximation, they concluded that the inclusion of space drastically affects the outcome of such biological interactions. Their results for the models with two species are reminiscent of those found in [15] for the spatially explicit Lotka–Volterra model and in [12] for a non-Mendelian diploid model: in the parameter region where the nonspatial model displays bistability, there is a dominant type that wins regardless of its initial density in the spatial model, while the inclusion of space also translates into a significant reduction of the parameter region where both species coexist. In the presence of three species, other disagreements between the spatial and nonspatial models appear. The main purpose of this paper is to give rigorous proofs of some of the results stated in [10] for the stochastic spatial model with two species, which we now describe in detail.

1.1. Model description

Each point of the d-dimensional square lattice is occupied by exactly one individual which is characterized by one of two possible types, say type 1 and type 2. Each individual is thought of as a producer and consumer of resources, and we describe the dynamics based on four parameters $a_{ij} \geq 0$ that denote the ability of an individual of type i to consume the resource produced by an individual of type j, which we call resource j. To avoid degenerate cases, we also assume that each resource can be consumed by at least one type: $a_{1j} + a_{2j} > 0$. Each individual dies at rate 1, i.e. the type at each vertex is updated at rate 1, with the new type being chosen at random from a certain neighborhood according to probabilities proportional to the ability of the neighbors to consume the resource produced by the individual to be updated. More formally, the model is a continuous-time Markov chain whose state at time t is a function

$$\eta_t : \mathbb{Z}^d \to \{1, 2\}$$

that maps the spatial structure into the set of types, thus representing the spatial configuration of the system at time t. The dynamics are described by the Markov generator

$$Lf(\eta) = \sum_{x \in \mathbb{Z}^d} \sum_{i \neq j} \frac{a_{ij} f_i(x, \eta)}{a_{1j} f_1(x, \eta) + a_{2j} f_2(x, \eta)} [f(\eta_{x,i}) - f(\eta)], \tag{1}$$

where $f_i(x, \eta)$ denotes the number of type-i neighbors

$$f_i(x, \eta) = \operatorname{card}\{y \neq x : |y_1 - x_1| + \cdots + |y_d - x_d| \leq M \text{ and } \eta(y) = i\}$$

and where the configuration $\eta_{x,i}$ is obtained from the configuration η by assigning type i to vertex x and leaving the type of all the other vertices unchanged. To avoid cumbersome notation, we will write later $x \sim y$ to indicate that vertices x and y are neighbors, i.e.

$$y \neq x \quad \text{and} \quad |y_1 - x_1| + |y_2 - x_2| + \cdots + |y_d - x_d| \leq M.$$

The positive integer M is referred to as the range of the interactions. The fraction in (1) means that the conditional probability that the new type is chosen to be type i given that vertex x is of type j at the time of an update is equal to the overall ability of the neighbors of type i to consume resource j divided by the overall ability of all the neighbors to consume resource j. Note that the assumption on the parameters, $a_{1j} + a_{2j} > 0$ for $j = 1, 2$, allows us to define

$$a_1 := a_{11}(a_{11} + a_{21})^{-1} \in [0, 1] \quad \text{and} \quad a_2 := a_{22}(a_{12} + a_{22})^{-1} \in [0, 1],$$

and to rewrite the Markov generator (1) in the form

$$Lf(\eta) = \sum_{x \in \mathbb{Z}^d} \sum_{i \neq j} \frac{(1 - a_j) f_i(x, \eta)}{a_j f_j(x, \eta) + (1 - a_j) f_i(x, \eta)} [f(\eta_{x,i}) - f(\eta)], \tag{2}$$

indicating that our model reduces to a two-parameter system. Note also that if one of the original parameters in (1) is equal to 0 then the denominator in

$$\frac{a_{ij} f_i(x, \eta)}{a_{1j} f_1(x, \eta) + a_{2j} f_2(x, \eta)} = \frac{(1 - a_j) f_i(x, \eta)}{a_j f_j(x, \eta) + (1 - a_j) f_i(x, \eta)}, \quad i \neq j, \tag{3}$$

may be equal to 0, in which case the numerator is equal to 0 as well. When this happens, we have the following alternatives and define (3) as follows.

Case 1: $a_j = f_i = 0$. In this case, motivated by the fact that there is no neighbor of type i, we define probability (3) at which vertex x becomes type i to be 0.

Case 2: $a_j = f_i = 1$. In this case, motivated by the fact that there are only neighbors of type i, we define probability (3) at which vertex x becomes type i to be 1.

We point out that, in addition to be the most natural from a biological point of view, the two assumptions above also make the transition probability (3) continuous with respect to the parameters, which is a key to some of our proofs.

1.2. The mean-field approximation

As pointed out in [10], one of the most interesting aspects of the spatial model is that it results in predictions that differ significantly from its nonspatial deterministic mean-field approximation, thus indicating that the spatial component plays a key role in these interactions. We recall that the mean-field model is obtained by assuming that the population is well mixing, and refer to Durrett and Levin [8] for more details. This assumption results in a system of coupled differential equations for the densities of each type, and since the densities sum to 1, the mean-field model of (2) reduces to the one-dimensional system

$$\frac{du_1}{dt} = \frac{a_{12}u_1u_2}{a_{12}u_1 + a_{22}u_2} - \frac{a_{21}u_2u_1}{a_{11}u_1 + a_{21}u_2}$$
$$= \frac{(1-a_2)u_1(1-u_1)}{a_2(1-u_1) + (1-a_2)u_1} - \frac{(1-a_1)u_1(1-u_1)}{a_1u_1 + (1-a_1)(1-u_1)}, \quad (4)$$

where u_i denotes the density of type-i individuals. Interestingly, the limiting behavior of the mean-field model depends on whether the parameters a_1 and a_2 are smaller or larger than one half, which has a biological interpretation in terms of altruism and selfishness. More precisely,

- type i is said to be *altruistic* whenever $a_i < \frac{1}{2}$ since in this case the resources it produces are more beneficial for the other type than for its own type,

- type i is said to be *selfish* whenever $a_i > \frac{1}{2}$ since in this case the resources it produces are more beneficial for its own type than for the other type.

The mean-field model has two trivial equilibria given by $e_1 = 1$ and $e_2 = 0$, respectively. By setting the right-hand side of (4) equal to 0 and assuming, in order to study the existence of nontrivial interior fixed points, that the product $u_1(1-u_1) \neq 0$, we find that

$$(1-a_2)\left(a_1 - \tfrac{1}{2}\right)u_1 = (1-a_1)\left(a_2 - \tfrac{1}{2}\right)(1-u_1),$$

from which it follows that

$$e_* = \frac{(1-a_1)(a_2 - 1/2)}{(1-a_2)(a_1 - 1/2) + (1-a_1)(a_2 - 1/2)}$$

is the unique fixed point that may differ from 0 and 1. To analyze the global stability of the fixed points, we observe that the sign of the derivative (4) is given by

$$\text{sgn}\left((1-a_2)\left(a_1 - \tfrac{1}{2}\right)u_1 - (1-a_1)\left(a_2 - \tfrac{1}{2}\right)(1-u_1)\right)$$
$$= \text{sgn}\left((1-a_2)\left(a_1 - \tfrac{1}{2}\right)(u_1 - e_*) + (1-a_1)\left(a_2 - \tfrac{1}{2}\right)(u_1 - e_*)\right), \quad (5)$$

which leads to the following three possible regimes for the mean-field model

Regime 1. *A selfish type always wins against an altruistic type.*

Proof. Without loss of generality, we may assume that $a_1 < \frac{1}{2} < a_2$, indicating that type 1 is altruistic whereas type 2 is selfish. In this case, the first line of (5) implies that the derivative

of u_1 is always negative when $u_1 \in (0, 1)$; hence, there is no fixed point in $(0, 1)$, and the trivial equilibrium e_1 is unstable while the trivial equilibrium e_2 is globally stable.

Regime 2. *In the presence of two selfish types, the system is bistable.*

Proof. Since $a_1 > \frac{1}{2}$ and $a_2 > \frac{1}{2}$, we have $e_* \in (0, 1)$. The second line of (5) implies that the derivative of u_1 has the same sign as $u_1 - e_*$, so the interior fixed point is unstable and the boundary equilibria locally stable: the system converges to e_1 if the initial density of type 1 is strictly larger than e_*, and to e_2 if the initial density of type 1 is strictly smaller than e_*.

Regime 3. *In the presence of two altruistic types, coexistence occurs.*

Proof. Since $a_1 < \frac{1}{2}$ and $a_2 < \frac{1}{2}$, the fixed point e_* again belongs to $(0, 1)$ as in the presence of two selfish types. However, the second line in (5) now implies that the derivative of u_1 has the same sign as $e_* - u_1$, so the interior fixed point is globally stable whereas the two boundary equilibria e_1 and e_2 are unstable.

1.3. The stochastic spatial model

In order to state our results for the spatial model, we start by defining some of the possible regimes the spin system (2) can exhibit.

Definition 1. For the stochastic spatial model (2), we say that

1. type i *wins* if, starting from any configuration with infinitely many type i,
$$\lim_{t \to \infty} \mathbb{P}(\eta_t(x) = i) = 1 \quad \text{for all } x \in \mathbb{Z}^d,$$

2. the system *clusters* if, starting from any translation invariant configuration,
$$\lim_{t \to \infty} \mathbb{P}(\eta_t(x) = \eta_t(y)) = 1 \quad \text{for all } x, y \in \mathbb{Z}^d,$$

3. type i *survives* if, starting from any configuration with infinitely many type i,
$$\liminf_{t \to \infty} \mathbb{P}(\eta_t(x) = i) > 0 \quad \text{for all } x \in \mathbb{Z}^d,$$

4. both types *coexist* if they both survive.

As pointed out above, the analytical results for the spatial and the nonspatial models differ in many aspects, which reveals the importance of the spatial component in the interactions. These disagreements are more pronounced in low dimensions and emerge when both types are selfish or both types are altruistic, while in the presence of one selfish type and one altruistic type the analytical results for both the spatial and nonspatial models agree.

To begin with, we focus on the one-dimensional nearest-neighbor process. In this case, one obtains a complete picture of the phase diagram based on a simple analysis of the interfaces of the process which almost evolve according to a system of annihilating random walks, whereas the more challenging study in higher dimensions and for larger ranges of interactions relies on very different techniques. We point out that the one-dimensional nearest-neighbor process appears as a degenerate case in which coexistence is not possible, while increasing the spatial dimension or the range of the interactions may result in coexistence of altruistic types.

Theorem 1. *Assume that $M = d = 1$. Then the following statements hold.*

1. *Except for the case $a_1 = a_2 = 1$, the process clusters.*

2. *Assuming, in addition, that $a_1 < a_2$, type 2 wins.*

The exclusion of the case $a_1 = a_2 = 1$ is simply due to the fact that, under this assumption, a vertex may flip to a new type if and only if all its neighbors already have this type, which implies that the process has infinitely many absorbing states. While, in the absence of space, one type wins regardless of the initial densities only in selfish-altruistic systems, Theorem 1 indicates that the most selfish/least altruistic type always wins in the one-dimensional nearest-neighbor case. In particular, coexistence no longer occurs even in altruistic systems. Although this case is pathological, it is also symptomatic of the general behavior of the spatial model for which the existence of a dominant type that always wins may occur even when both types are altruistic. The next result shows that, in any spatial dimension and for any dispersal range, selfish-altruistic interactions result in the same outcome in the spatial and nonspatial mean-field models.

Theorem 2. *Assume that $a_1 < \frac{1}{2} < a_2$. Then type 2 wins.*

Even though its rigorous proof is not straightforward, the result is predictable since, when one type is selfish and the other type is altruistic, regardless of the type of the vertex to be updated and the spatial configuration, each of the selfish neighbors is individually more likely to be selected to determine the new type than each of the altruistic neighbors. In contrast, in the case of selfish-selfish interactions covered by Theorem 3 below, the spatial and nonspatial models disagree in all dimensions and for all dispersal ranges. Recall that in this case the mean-field model is bistable, indicating that the long-term behavior is determined by both the parameters and the initial densities. The next result shows on the contrary that, including local interactions, even if both types are selfish, enough asymmetry in the parameters implies that the most selfish type wins.

Theorem 3. *For all $a_1 < 1$, there exists $\rho > 0$ such that type 2 wins when $a_2 > 1 - \rho$.*

Numerical simulations in two dimensions indicate more generally that, when type 2 is selfish, it wins whenever $a_1 < a_2$. In the neutral case, the obvious symmetry of the evolution rules implies that none of the types wins, but simulations indicate that the population again clusters, with boundaries getting sharper as the common value of $a_1 = a_2 > \frac{1}{2}$ increases. This suggests that, in the presence of two selfish types, the long-term behavior of the one-dimensional nearest-neighbor process occurs in two dimensions as well, a property that we believe is true in any spatial dimension and for any range of interactions. Looking finally at altruistic-altruistic interactions, recall that coexistence always occurs in the mean-field approximation whereas the least altruistic type always wins in the one-dimensional nearest-neighbor process. Our last two results show that the long-term behavior of the process in higher spatial dimensions and/or for larger dispersal ranges is intermediate between these two extreme behaviors.

Theorem 4. *For all $(M, d) \neq (1, 1)$, there exists $\rho > 0$ such that coexistence occurs when*

$$a_1 < \rho \quad and \quad a_2 < \rho.$$

In words, sufficiently strong mutual altruism translates into coexistence of both types. Numerical simulations of the two-dimensional system suggest in addition that two altruistic types coexist in the neutral case $a_1 = a_2 < \frac{1}{2}$; therefore, the coexistence region of the spatial model stretches up to the parameter point corresponding to voter model interactions. Interestingly, our last result shows that there exists a parameter region for which even altruistic types cannot coexist, thus indicating that the inclusion of local interactions translates into a reduction of the coexistence region in general and not only in the one-dimensional nearest-neighbor case.
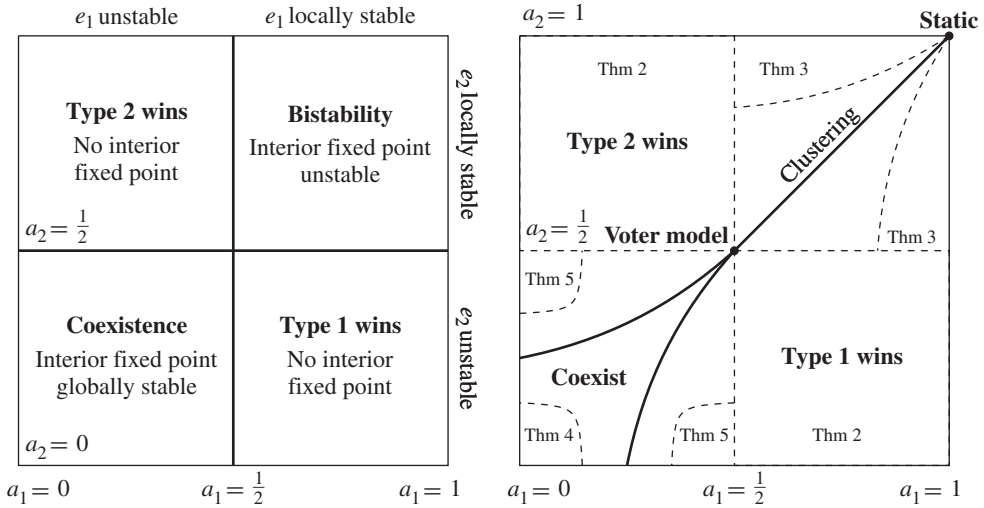
FIGURE 1: Phase diagrams of the nonspatial (*left*) and spatial (*right*) models.

Theorem 5. *There exists $\rho > 0$ such that type 2 wins whenever*

$$a_1 < \rho \quad and \quad a_2 > \tfrac{1}{2} - \rho.$$

We note that, even though the inclusion of local interactions always translates into a reduction of the coexistence region, numerical simulations indicate that this reduction becomes less pronounced while increasing the spatial dimension or the range of the interactions. The results for the mean-field model and spatial model when $(M, d) \neq (1, 1)$ are summarized in Figure 1. In the phase diagram of the spatial model, the dashed regions are those that are covered by our theorems, while the thick lines are the critical curves suggested by simulations. For a detailed analysis of the nonspatial model and additional results based on numerical simulations for the spatial model in the presence of three species, we refer the reader to the companion paper [10].

The rest of this paper is devoted to the proofs. To avoid cumbersome notation and messy calculations, we prove all our results, with the exception of Theorem 4, when the range of the interactions is equal to 1. However, we point out that all the techniques used in the proofs easily extend to larger dispersal ranges. The reason for excluding Theorem 4 from this rule is that it is the only result that shows a disagreement between the one-dimensional nearest-neighbor case and the other cases, thus indicating that the range of the interactions plays a specific role in the presence of strong altruistic-altruistic interactions.

2. The one-dimensional spatial model

This section is devoted to the proof of Theorem 1 which gives a complete picture of the long-term behavior of the one-dimensional process with nearest-neighbor interactions. To begin with, we study the evolution in the nonneutral case, and look at the process starting with only type-2 individuals in an interval of length $N + 1$. The following lemma shows that the survival probability of the individuals of type 1 decreases exponentially with N when $a_1 < a_2$.

Lemma 1. *Assume that $M = d = 1$ and $a_1 < a_2$. Then*

$$\mathbb{P}(\eta_t(x) \to 2 \text{ for all } x \in \mathbb{Z} \mid \eta_0(x) = 2 \text{ for all } x \in [0, N] \cap \mathbb{Z}) \geq 1 - c_0^N,$$

where $c_0 := (1 - a_2)(1 - a_1)^{-1} < 1$.

Proof. First, we observe that the process is attractive, which directly follows from the monotonicity of the probabilities in (3) with respect to the fractions of neighbors of each type. In particular, it suffices to prove the result when starting with exactly $N + 1$ individuals of type 2. To understand how the number of type-2 vertices evolves, let $x \in \mathbb{Z}$ and assume that the neighbors of x have different types. Then, the rate at which vertex x jumps from type 2 to type 1 is

$$c_- := \frac{(1 - a_2) f_1(x, \eta)}{(1 - a_2) f_1(x, \eta) + a_2 f_2(x, \eta)} = \frac{1 - a_2}{(1 - a_2) + a_2} = 1 - a_2, \quad (6)$$

while the rate at which it jumps from type 1 to type 2 is

$$c_+ := \frac{(1 - a_1) f_2(x, \eta)}{a_1 f_1(x, \eta) + (1 - a_1) f_2(x, \eta)} = \frac{1 - a_1}{a_1 + (1 - a_1)} = 1 - a_1. \quad (7)$$

In particular, as long as there are at least two individuals of type 2, the number of such individuals performs a random walk that jumps to the right at rate $2c_+$ and to the left at rate $2c_-$ since, due to one-dimensional nearest-neighbor interactions and the choice of the initial configuration, at all times the set of vertices of type 2 is an interval. Moreover, on the event that type 2 survives, the leftmost vertex of type 2 converges almost surely to $-\infty$, while the rightmost vertex of type 2 converges almost surely to $+\infty$; therefore, the probability to be estimated is equal to the probability that the asymmetric random walk described above tends to ∞. The result then reduces to a standard estimate for asymmetric random walks. More precisely, let Z_n denote the number of individuals of type 2 after n updates of the system, and let

$$P_i := \mathbb{P}(Z_n = 1 \text{ for some } n \geq 0 \mid Z_0 = i).$$

Then, using a first-step analysis, for all $i \geq 2$ we have

$$(c_+ + c_-) P_i = c_+ P_{i+1} + c_- P_{i-1}$$
$$\text{and} \quad P_{i+1} - P_i = c_0 (P_i - P_{i-1}) = \cdots = c_0^{i-1} (P_2 - P_1),$$

where $c_0 = c_-/c_+ = (1 - a_2)(1 - a_1)^{-1}$. Since $P_1 = 1$, we deduce that $P_i = c_0^{i-1}$ and

$$\mathbb{P}(\eta_t(x) \to 2 \text{ for all } x \in \mathbb{Z} \mid \eta_0(x) = 2 \text{ for all } x \in [0, N] \cap \mathbb{Z})$$
$$\geq \mathbb{P}(\eta_t(x) \to 2 \text{ for all } x \in \mathbb{Z} \mid \{x : \eta_0(x) = 2\} = [0, N] \cap \mathbb{Z})$$
$$\geq 1 - P_{N+1}$$
$$= 1 - c_0^N.$$

This completes the proof.

With Lemma 1 in hand, we can now establish the first part of Theorem 1 in the nonneutral case as well as the second part of the theorem.

Lemma 2. *Assume that $M = d = 1$ and $a_1 < a_2$. Then type 2 wins.*

Proof. Let N be a positive integer. Then, starting from any initial configuration with infinitely many vertices of type 2, with probability 1, there exist $z \in \mathbb{Z}$ and $t < \infty$ such that

$$\eta_t(x) = 2 \quad \text{for } x = z, z + 1, \ldots, z + N.$$

Using, in addition, the facts that the process is attractive, and the evolution rules are translation invariant in space and time, Lemma 1 implies that

$$\mathbb{P}(\eta_t(x) \to 2 \text{ for all } x \in \mathbb{Z} \mid \text{card}\{x : \eta_0(x) = 2\} = \infty)$$
$$\geq \mathbb{P}(\eta_t(x) \to 2 \text{ for all } x \in \mathbb{Z} \mid \eta_0(x) = 2 \text{ for all } x \in [0, N] \cap \mathbb{Z})$$
$$\geq 1 - c_0^N.$$

Since this holds for all N and since $c_0 < 1$, the lemma follows.

To complete the proof of Theorem 1, it remains to show that, when $a_1 = a_2 \neq 1$, the process clusters. Due to the particular geometry of the configurations in the one-dimensional nearest-neighbor case, the proof reduces to the analysis of an auxiliary process that we will call the interface process: the spin system (ξ_t) defined on the translated lattice $\mathbb{D} = \mathbb{Z} + \frac{1}{2}$ by setting

$$\xi_t(v) = \left| \eta_t\left(v + \tfrac{1}{2}\right) - \eta_t\left(v - \tfrac{1}{2}\right) \right| \quad \text{for all } v \in \mathbb{D}.$$

In words, the process has a 1 at sites $v \in \mathbb{D}$ which are located between two vertices that have different types, and a 0 at sites which are located between two vertices that have the same type, so the process keeps track of the interfaces of the spatial model.

Lemma 3. *Assume that $M = d = 1$ and $a_1 = a_2 \neq 1$. Then the process clusters.*

Proof. Thinking of each site $v \in \mathbb{D}$ as being occupied by a particle if $\xi(v) = 1$ and as empty otherwise, the idea of the proof is to establish almost-sure extinction of this system of particles, which is equivalent to clustering of the spatial model. First, the proof of Lemma 1 indicates that a vertex of either type, say type i, changes its type

- at rate 1 if none of its two nearest neighbors is of type i,
- at rate $c_- = c_+ > 0$ defined in (6) and (7) if its neighbors have different types.

This induces the following dynamics for the interface process: a particle at site $v \in \mathbb{D}$ jumps to each of its empty nearest neighbors at rate c_+ and two particles a distance 1 apart annihilate each other at rate 1. More formally, the Markov generator is given by

$$L_\xi f(\xi) = \sum_{v \in \mathbb{D}} \sum_{|w-v|=1} c_+ \mathbf{1}\{\xi(v) \neq \xi(w)\}[f(\xi_{v \leftrightarrow w}) - f(\xi)]$$
$$+ \sum_{v \in \mathbb{D}} \sum_{|w-v|=1} \mathbf{1}\{\xi(v) = \xi(w)\}[f(\xi_{v,w}) - f(\xi)],$$

where configuration $\xi_{v \leftrightarrow w}$ is obtained from ξ by exchanging the contents of vertices v and w while configuration $\xi_{v,w}$ is obtained by killing the particles at v and w if they exist. Since particles can only annihilate, the probability that a given site is occupied by a particle decreases in time so it has a limit when time goes to ∞. Since, in addition, the evolution rules of the process are translation invariant, this limit does not depend on the site under consideration:

$$\text{there exists } a \text{ such that } \lim_{t \to \infty} \mathbb{P}(\xi_t(v) = 1) = a \text{ for all } v \in \mathbb{D}.$$

In other respects, given any two particles alive at time t, at some random time $s > t$ almost surely finite, either one of these particles is killed due to a collision with a third particle or

both particles annihilate each other. The latter follows immediately from the fact that one-dimensional random walks are recurrent, while each time two particles are a distance 1 apart, there is a positive probability that they annihilate at their next jump when $a_1 = a_2 \neq 1$. This implies that the limit a must equal 0, from which we conclude that

$$\lim_{t \to \infty} \mathbb{P}(\eta_t(x) \neq \eta_t(y)) \leq \lim_{t \to \infty} \sum_{i=1}^{y-x} \mathbb{P}\left(\xi_t\left(x + i - \frac{1}{2}\right) = 1\right) = 0$$

for all $x, y \in \mathbb{Z}$. This completes the proof.

3. Altruistic-selfish interactions

This section is devoted to the proof of Theorem 2 which states that, in the presence of altruistic-selfish interactions, the selfish type always wins. The proof is divided into four steps. The first step is to show that, under the assumptions of the theorem, the set of type 2 dominates its counterpart in a process that we will call perturbation of the voter model. In particular, in order to establish the theorem, it suffices to prove that type 2 wins for this new process. The reason for introducing a perturbation of the voter model is that, contrary to process (2), it can be studied using duality techniques, and the second step of the proof is to describe its dual process in detail, while the third step is to construct selected dual paths that are key to proving that type 2 survives. Finally, the fourth step combines these selected dual paths to a block construction in order to prove that not only type 2 survives but also type 1 goes extinct. Before giving the details of the proof, we note that the third step will be used again in the proof of Theorem 5 while the fourth step will be used again in the proofs of both Theorems 3 and 5, but these two steps are detailed in this section only. We also point out that Theorem 2 can be proved without the use of a block construction, but since it is needed in the proofs of Theorems 3 and 5, we follow the same approach for all three theorems.

3.1. Coupling with a voter model perturbation

We first observe that, under the assumptions of the theorem, there exists a constant $\rho > 0$ fixed from now on such that

$$a_1 < \frac{1 - \rho}{2} \quad \text{and} \quad a_2 > \frac{1 + \rho}{2}, \tag{8}$$

in which case the set of type 2 for the spatial model (2) dominates stochastically its counterpart in a certain perturbation of the voter model that we denote later by (ξ_t). The dynamics of this voter model perturbation depend on the single parameter

$$\varepsilon := \frac{\rho}{d(1 - \rho) + \rho} \tag{9}$$

and can be described as follows. As in the original spatial model, the type at each vertex x is updated at rate 1, but the new type is now chosen according to the following rules:

- with probability $1 - \varepsilon > 0$, the new type at vertex x is chosen uniformly at random from the set of the nearest neighbors,

- with the residual probability $\varepsilon > 0$, the new type at vertex x is chosen to be type 1 if all the nearest neighbors are of type 1, and type 2 otherwise.

More formally, the dynamics are described by the Markov generator

$$L_\xi f(\xi) = \sum_{x \in \mathbb{Z}^d} [(1 - \varepsilon) f_1(x, \xi) + \varepsilon \mathbf{1}\{f_2(x, \xi) = 0\}][f(\xi_{x,1}) - f(\xi)]$$

$$+ \sum_{x \in \mathbb{Z}^d} [(1 - \varepsilon) f_2(x, \xi) + \varepsilon \mathbf{1}\{f_2(x, \xi) \neq 0\}][f(\xi_{x,2}) - f(\xi)], \qquad (10)$$

where configuration $\xi_{x,i}$ is obtained from ξ by assigning type i to vertex x and leaving the type of all the other vertices unchanged. Then we have the following result.

Lemma 4. *Assume that $\eta_0 \equiv \xi_0$ and (8) holds. Then*

$$\mathbb{P}(\eta_t(x) = 2) \geq \mathbb{P}(\xi_t(x) = 2) \quad \text{for all } x \in \mathbb{Z}^d \text{ and all } t > 0.$$

Proof. This follows from certain inequalities between the transition rates. When all the neighbors of vertex x are of the same type at the time of an update, then vertex x becomes of this type with probability 1 in both processes. Also, given that x is of type 2 and has at least one neighbor of each type, the rate at which it becomes 1 for process (2) is bounded from above by

$$\frac{(1 - a_2) f_1(x, \eta)}{(1 - a_2) f_1(x, \eta) + a_2 f_2(x, \eta)} \leq \frac{(1 - \rho) f_1(x, \eta)}{(1 - \rho) f_1(x, \eta) + (1 + \rho) f_2(x, \eta)}$$

$$= \frac{(1 - \rho) f_1(x, \eta)}{(1 - \rho) + 2\rho f_2(x, \eta)}$$

$$\leq \frac{d(1 - \rho)}{d(1 - \rho) + \rho} f_1(x, \eta)$$

$$= (1 - \varepsilon) f_1(x, \eta), \qquad (11)$$

which is the rate at which it becomes 1 for process (10). Similarly, given that x is of type 1 and has at least one neighbor of each type, the rate at which it becomes 2 for process (2) is

$$\frac{(1 - a_1) f_2(x, \eta)}{a_1 f_1(x, \eta) + (1 - a_1) f_2(x, \eta)} \geq \frac{(1 + \rho) f_2(x, \eta)}{(1 - \rho) f_1(x, \eta) + (1 + \rho) f_2(x, \eta)}$$

$$= \frac{(1 + \rho) f_2(x, \eta)}{(1 - \rho) + 2\rho f_2(x, \eta)}$$

$$\geq \frac{d(1 - \rho) f_2(x, \eta) + \rho}{d(1 - \rho) + \rho}$$

$$= (1 - \varepsilon) f_2(x, \eta) + \varepsilon, \qquad (12)$$

which is the rate at which it becomes 2 for process (10). To see that the last inequality is indeed true, we introduce the functions

$$g(z) = \frac{(1 + \rho)z}{(1 - \rho) + 2\rho z} \quad \text{and} \quad h(z) = \frac{d(1 - \rho)z + \rho}{d(1 - \rho) + \rho},$$

and observe that we have the equalities

$$g(1) = h(1) = 1 \quad \text{and} \quad g((2d)^{-1}) = h((2d)^{-1}) = \frac{1 + \rho}{2d(1 - \rho) + 2\rho}.$$

Since, in addition, g is concave on $(0, 1)$ whereas h is linear, the last inequality in (12) follows. The inequalities on the transition rates in (11) and (12) give the lemma.

3.2. Duality with branching coalescing random walks

In view of Lemma 4, standard coupling arguments—see [13, Section II.1] for details about coupling techniques—imply that, to prove the theorem, it suffices to prove that type 2 wins for process (10). Recall that the reason for introducing this process is that, contrary to the original spatial model, it can be studied using duality techniques. First, we construct process (10) graphically as follows:

- for each oriented edge $e = (x, y) \in \mathbb{Z}^d \times \mathbb{Z}^d$ with $x \sim y$,

  ○ we let $\Lambda(x, y)$ be a Poisson point process with intensity $(2d)^{-1}(1 - \varepsilon)$,

  ○ we draw an arrow from vertex y to vertex x at times $t \in \Lambda(x, y)$ to indicate that the individual at x mimics the individual at y,

- for each vertex $x \in \mathbb{Z}^d$,

  ○ we let $\Delta(x)$ be a Poisson point process with intensity ε,

  ○ we draw a set of $2d$ arrows from each $y \sim x$ to vertex x at times $t \in \Delta(x)$ to indicate that x becomes of type 1 if all its neighbors are of type 1, and of type 2 otherwise.

We refer the reader to the left-hand diagram of Figure 2 for an illustration of the graphical representation. The type at any space–time point can be deduced from the graphical representation and the configuration of the system at earlier times by going backwards in time. Declare that a dual path from space–time point (x, T) down to $(y, T - s)$ exists, which we write as $(x, T) \downarrow (y, T - s)$, whenever there are sequences of times and vertices

$$s_0 = T - s < s_1 < \cdots < s_{m+1} = T \quad \text{and} \quad x_0 = y, x_1, \ldots, x_m = x$$
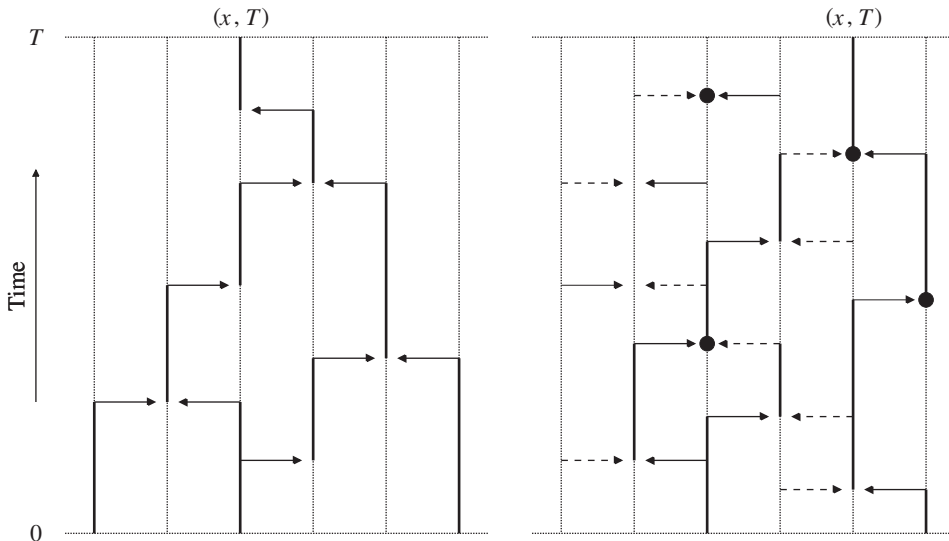

FIGURE 2: Pictures related to the proofs of Theorems 2 (*left*) and 5 (*right*).

such that the following two conditions hold:

- for $i = 1, 2, \ldots, m$, there is an arrow from x_{i-1} to x_i at time s_i, and

- for $i = 0, 1, \ldots, m$, there is no arrow that points at the segments $\{x_i\} \times (s_i, s_{i+1})$.

The dynamics imply that vertex x is of type 1 at time T if and only if

$$\xi_0(y) = 1 \quad \text{for all } y \in \mathbb{Z}^d \text{ such that } (x, T) \downarrow (y, 0). \tag{13}$$

Note that the set-valued process

$$\hat{\xi}_s(x, T) = \{y \in \mathbb{Z}^d : (x, T) \downarrow (y, T - s)\} \quad \text{for all } 0 \le s \le T$$

defines a system of branching coalescing random walks where particles jump at rate $1 - \varepsilon$ to one of their nearest neighbors chosen uniformly at random and split at rate ε into $2d$ particles which are sent to each of their neighbors. In addition, whenever two particles are located on the same vertex, they coalesce. See the left-hand diagram of Figure 2 for an illustration, where the system of random walks is represented by thick lines. This process is reminiscent, though not identical, of the dual process of the biased voter model [4], and we also refer the reader to [1] and [11] for similar dual processes.

3.3. Selected dual paths and random walk estimates

To bound the probability that a given space–time point (x, T) is of type 1, we now construct a dual path $(X_t(x))$ that keeps track of a specific particle in the dual process. For each $i = 1, 2, \ldots, d$, we define the hyperplane

$$H_i := \{z = (z_1, z_2, \ldots, z_d) \in \mathbb{R}^d : z_i = 0\}$$

as well as the deterministic times $T_i = i c_1 N$, where $c_1 = 4\varepsilon^{-1}$. Recall that the constant ε is defined in (9). Also, we let $\text{dist}(x, H)$ denote the Euclidean distance between a vertex x and its orthogonal projection on a set H. The dual path starts at $X_0(x) = x$ and is defined recursively from the graphical representation as follows. For all $t > 0$, define

$$s(t) = \inf\{s > t : T - s \in \Delta(X_t(x)) \text{ or } T - s \in \Lambda(X_t(x), y) \text{ for some } y \sim X_t(x)\}.$$

For all $s \in (t, s(t))$, we set $X_s(x) = X_t(x)$, while at time $s(t)$ we have the following alternatives.

1. If $T - s(t) \in \Lambda(X_t(x), y)$ for some $y \sim X_t(x)$ then $X_{s(t)}(x) = y$.

2. If $T - s(t) \in \Delta(X_t(x))$ and $t \in (T_{i-1}, T_i)$ for some $i = 1, 2, \ldots, d$, then

$$X_{s(t)}(x) \sim X_t(x) \quad \text{and} \quad \text{dist}(X_{s(t)}(x), H_i) < \text{dist}(X_t(x), H_i).$$

3. If $T - s(t) \in \Delta(X_t(x))$ and $t > T_d$ then

$$\mathbb{P}(X_{s(t)}(x) = y) = (2d)^{-1} \quad \text{for all } y \sim X_t(x).$$

Note that there is a unique $X_{s(t)}(x)$ that satisfies condition 2 above, so the dual path is well defined. In words, the dual path travels backwards in time down the graphical representation following the arrows from tip to tail. When a set of $2d$ arrows is encountered, the process jumps in the direction that makes it closer to the hyperplane H_i between time T_{i-1} and time T_i, while

after the last time T_d the process jumps uniformly at random in all directions. This, together with the properties of the graphical representation introduced above, implies that the selected dual path always jumps at rate 1. At each jump, with probability $1 - \varepsilon$, the target site is chosen uniformly at random from the set of the nearest neighbors, whereas, with probability ε, the process moves so as to decrease its distance to H_1 until time T_1 when a similar mechanism operates in the direction of the second axis, and so on. In order to later compare process (10) with oriented site percolation, we start by collecting important properties of the selected dual path. These properties are based on random walk estimates and are given in the following three lemmas.

Lemma 5. *There exist $C_1 < \infty$ and $\gamma_1 > 0$ such that, for all $i = 1, 2, \ldots, d$,*

$$\mathbb{P}\big(\mathrm{dist}(X_{T_i}(x), H_i) > \tfrac{1}{2}N \mid \mathrm{dist}(X_{T_{i-1}}(x), H_i) \leq 4N\big) \leq C_1 \exp(-\gamma_1 N)$$

for all sufficiently large N.

Proof. First, we observe that, for all $t \in (T_{i-1}, T_i)$, we have

$$\mathrm{dist}(X_t(x), H_i) \to \begin{cases} \mathrm{dist}(X_t(x), H_i) + 1 & \text{at rate } (2d)^{-1}(1 - \varepsilon), \\ \mathrm{dist}(X_t(x), H_i) - 1 & \text{at rate } (2d)^{-1}(1 - \varepsilon) + \varepsilon, \end{cases}$$

from which we deduce that, for all $t \in (T_{i-1}, T_i)$,

$$\lim_{h \to 0} h^{-1} \mathbb{E}(\mathrm{dist}(X_{t+h}(x), H_i) - \mathrm{dist}(X_t(x), H_i)) = -\varepsilon. \tag{14}$$

In particular, recalling the definitions of T_i and T_{i-1}, we obtain

$$\mathbb{E}(\mathrm{dist}(X_{T_i}(x), H_i) \mid \mathrm{dist}(X_{T_{i-1}}(x), H_i) = 4N) = 4N - \varepsilon(T_i - T_{i-1}) = 0.$$

The lemma then follows from large deviation estimates for the Poisson distribution and the binomial distribution together with standard coupling arguments.

Lemma 6. *For any constant $C > 4d\varepsilon^{-1}$ and all $i = 1, 2, \ldots, d$,*

$$\mathbb{P}\big(\mathrm{dist}(X_t(x), H_i) > N \text{ for some } t \in (T_i, CN) \mid \mathrm{dist}(X_{T_i}(x), H_i) \leq \tfrac{1}{2}N\big) \leq 3 \exp(-\sqrt{N})$$

for all sufficiently large N.

Proof. Denote by (Z_n) the discrete-time, one-dimensional simple symmetric random walk, and introduce the discrete random variable J_s that counts the number of times the ith coordinate of the process $(X_t(x))$ jumps by time s. Since the ith coordinate evolves according to the continuous-time simple symmetric random walk run at rate $d^{-1}(1 - \varepsilon)$ after time T_i, we have

$$\mathbb{P}\big(\mathrm{dist}(X_t(x), H_i) > N \text{ for some } t \in (T_i, CN) \mid \mathrm{dist}(X_{T_i}(x), H_i) \leq \tfrac{1}{2}N\big)$$
$$\leq \mathbb{P}\big(J_{CN} > N_0\big) + \mathbb{P}\big(Z_n \notin (-N, N) \text{ for some } n \leq N_0 \mid Z_0 \in \big(-\tfrac{1}{2}N, \tfrac{1}{2}N\big)\big)$$

for all $N_0 > 0$. Taking $N_0 = d^{-1}CN$, large deviation estimates for the Poisson distribution imply that the first term on the right-hand side can be bounded by

$$\mathbb{P}(J_{CN} > N_0) \leq \mathbb{P}(J_{CN} > d^{-1}(1 - \varepsilon)CN) \leq C_2 \exp(-\gamma_2 N) \leq \exp(-\sqrt{N})$$

for suitable constants $C_2 < \infty$ and $\gamma_2 > 0$, and all sufficiently large N. To estimate the second term, we use the reflection principle and Chebyshev's inequality. For all $\theta > 0$, we have

$$
\begin{aligned}
\mathbb{P}\big(Z_n &\notin (-N, N) \text{ for some } n \le N_0 \mid Z_0 \in \big(-\tfrac{1}{2}N, \tfrac{1}{2}N\big)\big) \\
&\le \mathbb{P}\big(Z_n \notin \big(-\tfrac{1}{2}N, \tfrac{1}{2}N\big) \text{ for some } n \le N_0 \mid Z_0 = 0\big) \\
&\le 2\mathbb{P}\big(Z_{N_0} \notin \big(-\tfrac{1}{2}N, \tfrac{1}{2}N\big) \mid Z_0 = 0\big) \\
&\le 2\exp\big(-\tfrac{1}{2}\theta N\big)\mathbb{E}(\exp(\theta Z_{N_0}) \mid Z_0 = 0) \\
&\le 2\exp\big(-\tfrac{1}{2}\theta N\big)[\mathbb{E}(\exp(\theta Z_1) \mid Z_0 = 0)]^{N_0} \\
&\le 2\exp\big(-\tfrac{1}{2}\theta N + N_0 \ln \phi(\theta)\big),
\end{aligned}
$$

where the function $\phi(\theta)$ is the moment generating function of Z_1. Since

$$
\ln \phi(\theta) = \ln\left(\frac{e^\theta + e^{-\theta}}{2}\right) = \ln\left(1 + \frac{\theta^2}{2} + o(\theta^2)\right) \le \theta^2
$$

when $\theta > 0$ is small enough, taking $\theta = 4/\sqrt{N}$, we can conclude that

$$
\begin{aligned}
\mathbb{P}\big(Z_n &\notin (-N, N) \text{ for some } n \le N_0 \mid Z_0 \in \big(-\tfrac{1}{2}N, \tfrac{1}{2}N\big)\big) \\
&\le 2\exp\big(-\tfrac{1}{2}\theta N + N_0 \ln \phi(\theta)\big) \\
&\le 2\exp(-2\sqrt{N} + 16d^{-1}C) \\
&\le 2\exp(-\sqrt{N})
\end{aligned}
$$

for all sufficiently large N. This completes the proof.

Lemma 7. *Let $C_3 > 4d\varepsilon^{-1}$ and $x \in (-2N, 2N]^d$. Then, for all large N,*

$$
\mathbb{P}(X_t(x) \notin (-4N, 4N]^d \text{ for some } t < C_3N \text{ or } X_{C_3N}(x) \notin (-N, N]^d) \le 7d\exp(-\sqrt{N}).
$$

Proof. Since $\mathrm{dist}(X_t(x), H_i)$ is stochastically smaller than the rate 1 simple symmetric random walk on the set of nonnegative integers with a reflecting boundary at 0 and starting from the same initial state, the proof of Lemma 6 directly implies that

$$
\mathbb{P}(\mathrm{dist}(X_t(x), H_i) > 4N \text{ for some } t < C_3N) \le 3\exp(-\sqrt{4N}) \tag{15}
$$

for all large N and all $i = 1, 2, \ldots, d$. Moreover, by Lemmas 5 and 6,

$$
\begin{aligned}
\mathbb{P}(\mathrm{dist}&(X_{C_3N}(x), H_i) > N \mid \mathrm{dist}(X_t(x), H_i) \le 4N \text{ for all } t < C_3N) \\
&\le \mathbb{P}\big(\mathrm{dist}(X_{T_i}(x), H_i) > \tfrac{1}{2}N \mid \mathrm{dist}(X_{T_{i-1}}(x), H_i) \le 4N\big) \\
&\quad + \mathbb{P}\big(\mathrm{dist}(X_t(x), H_i) > N \text{ for some } t \in (T_i, C_3N) \mid \mathrm{dist}(X_{T_i}(x), H_i) \le \tfrac{1}{2}N\big) \\
&\le C_1 \exp(-\gamma_1 N) + 3\exp(-\sqrt{N}) \tag{16}
\end{aligned}
$$

for all large N and all $i = 1, 2, \ldots, d$. Finally, combining (15) and (16), we conclude that

$$
\begin{aligned}
\mathbb{P}(X_t(x) &\notin (-4N, 4N]^d \text{ for some } t < C_3N \text{ or } X_{C_3N}(x) \notin (-N, N]^d) \\
&\le \mathbb{P}(X_t(x) \notin (-4N, 4N]^d \text{ for some } t < C_3N) \\
&\quad + \mathbb{P}(X_{C_3N}(x) \notin (-N, N]^d \text{ and } X_t(x) \in (-4N, 4N]^d \text{ for all } t < C_3N) \\
&\le 3d\exp(-\sqrt{4N}) + dC_1 \exp(-\gamma_1 N) + 3d\exp(-\sqrt{N}) \\
&\le 7d\exp(-\sqrt{N})
\end{aligned}
$$

for all sufficiently large N. This completes the proof.

3.4. Block construction

To complete the proof of Theorem 2, the last step is to use a rescaling argument, a technique also known as block construction, which has been introduced by Bramson and Durrett [3]. The general idea of the block construction is to couple certain good events related to the interacting particle system properly rescaled in space and time with the set of open sites of an oriented site percolation process on the lattice

$$\mathcal{H} = \{(z, n) = (z_1, \ldots, z_d, n) \in \mathbb{Z}^d \times \mathbb{Z} : z_1 + \cdots + z_d + n \text{ is even and } n \geq 0\}.$$

We refer the reader to Section 4 of [7] for more details about this technique and a rigorous definition of oriented site percolation. Lemmas 5–7 give us the appropriate space and time scales to compare process (10) with oriented percolation. More precisely, we let $T = C_3 N$, where the constant C_3 is as in Lemma 7, and declare $(z, n) \in \mathcal{H}$ to be good whenever the event

$$\Omega(z, n) = \{\xi_{nT}(x) = 2 \text{ for all } x \in Nz + (-N, N]^d \cap \mathbb{Z}^d\}$$

occurs. Then we have the following lemma.

Lemma 8. *For all sufficiently large N,*

$$\mathbb{P}(X_t(x) \notin (-4N, 4N]^d \text{ for some } (x, t) \in (-2N, 2N]^d \times (0, T))$$
$$+ \mathbb{P}((\pm e_i, 1) \text{ is not good for some } i = 1, 2, \ldots, d \mid (0, 0) \text{ is good})$$
$$\leq 7d((4N)^d + 1) \exp(-\sqrt{N}),$$

where e_i denotes the ith d-dimensional unit vector.

Proof. By construction, there exists a dual path $(x, T) \downarrow (X_T(x), 0)$. Therefore, Lemma 7 and the duality relationship (13) imply that

$$\mathbb{P}((\pm e_i, 1) \text{ is not good for some } i = 1, 2, \ldots, d \mid (0, 0) \text{ is good})$$
$$= \mathbb{P}(\xi_T(x) = 1 \text{ for some } x \in (-2N, 2N]^d \mid \xi_0(y) = 2 \text{ for all } y \in (-N, N]^d)$$
$$\leq \mathbb{P}(\xi_0(X_T(x)) = 1 \text{ for some } x \in (-2N, 2N]^d \mid \xi_0(y) = 2 \text{ for all } y \in (-N, N]^d)$$
$$\leq (4N)^d \sup_{x \in (-2N, 2N]^d} \mathbb{P}(\xi_0(X_T(x)) = 1 \mid \xi_0(y) = 2 \text{ for all } y \in (-N, N]^d)$$
$$\leq (4N)^d \sup_{x \in (-2N, 2N]^d} \mathbb{P}(X_T(x) \notin (-N, N]^d)$$
$$\leq 7d(4N)^d \exp(-\sqrt{N})$$

for all large N. Lemma 7 also implies that, for all $x \in (-2N, 2N]^d$, we have

$$\mathbb{P}(X_t(x) \notin (-4N, 4N]^d \text{ for some } t \in (0, T)) \leq 7d \exp(-\sqrt{N})$$

for all sufficiently large N. The lemma follows.

The right-hand side of the inequality in the statement of Lemma 8 tends to 0; therefore, both probabilities on the left-hand side tend to 0, as N tends to ∞. Convergence to 0 of the second probability together with the translation invariance of the evolution rules of the process in space and time imply that, for all $\gamma > 0$, there exists N large such that

$$\mathbb{P}((z \pm e_i, n + 1) \text{ is good for all } i = 1, 2, \ldots, d \mid (z, n) \text{ is good}) \geq 1 - \gamma. \tag{17}$$

Convergence to 0 of the first probability in the statement of the lemma implies that, with probability close to 1 when N is large, the good event in (17) depends only on the realization

of the graphical representation in a finite space–time box which, together with (17), implies the existence of events $G_{z,n}$ such that

1. $G_{z,n}$ is measurable with respect to the graphical representation in

$$\{Nz + (-4N, 4N]^d\} \times [nT, (n+1)T] \quad \text{for all } (z, n) \in \mathcal{H},$$

2. for all $\gamma > 0$, there exists N large such that $\mathbb{P}(G_{z,n}) \geq 1 - \gamma$,

3. we have the inclusions of events

$$G_{z,n} \cap \Omega(z, n) \subset \Omega(z \pm e_i, n+1) \quad \text{for all } i = 1, 2, \ldots, d.$$

These three conditions are the assumptions of Theorem 4.3 of [7], from which we deduce the existence of a coupling such that the set of good sites contains the set of wet sites of a certain oriented site percolation process with parameter $1 - \gamma$ and finite range of dependency when the scale parameter N is large. Since such a percolation process is supercritical when γ is small enough, starting with infinitely many vertices of type 2, we have

$$\liminf_{t \to \infty} \mathbb{P}(\xi_t(x) = 2) > 0 \quad \text{for all } x \in \mathbb{Z}^d.$$

This only proves survival of the type-2 individuals, and the last step of the proof is to also establish extinction of the type-1 individuals, which is not obvious since oriented site percolation has a positive density of closed sites. The existence of an in-all-directions expanding region void of type 1 is proved in the following lemma which we will apply again to prove Theorems 3 and 5.

Lemma 9. *For all $x \in \mathbb{Z}^d$, we have $\mathbb{P}(\xi_t(x) = 1) \to 0$ as time $t \to \infty$.*

Proof. The proof follows an idea of Durrett [6] which relies on the lack of percolation of the dry sites, where a site in the percolation process is said to be dry if it is not wet. First, we define the oriented graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ with common vertex set $\mathcal{H}$ and respective edge sets

$$E_1 = \{((z_1, n_1), (z_2, n_2)) \in \mathcal{H} \times \mathcal{H} : z_1 - z_2 \in \{\pm e_1, \ldots, \pm e_d\} \text{ and } n_1 = n_2 - 1\},$$
$$E_2 = E_1 \cup \{((z_1, n_1), (z_2, n_2)) \in \mathcal{H} \times \mathcal{H} : z_1 - z_2 \in \{\pm 2e_1, \ldots, \pm 2e_d\} \text{ and } n_1 = n_2\}.$$

Given a realization of the percolation process in which sites are open with probability $1 - \gamma$, we say that there is a regular dry path from $(0, 0)$ to (z, n) if there is a sequence

$$(z_0, n_0) = (0, 0), (z_1, n_1), \ldots, (z_k, n_k) = (z, n) \in \mathcal{H}$$

such that the following two conditions hold:

1. $((z_i, n_i), (z_{i+1}, n_{i+1})) \in E_1$ for all $i = 0, 1, \ldots, k - 1$, and

2. the site (z_i, n_i) is dry for all $i = 0, 1, \ldots, k$.

Extending condition 1 by replacing E_1 with the larger edge set E_2, we call the sequence a generalized dry path. Assuming that initially only site $(0, 0)$ is closed, we introduce

- $D_0(\mathcal{G}_1)$, the set of sites that can be reached from the origin by a regular dry path, and

- $D_0(\mathcal{G}_2)$, the set of sites that can be reached from the origin by a generalized dry path.

We call these sets the regular dry cluster and the generalized dry cluster, respectively. Returning to the particle system, in order to prove extinction of type 1, we first modify our definition of good sites: a site $(z, n) \in \mathcal{H}$ is now said to be good if the event

$$\bar{\Omega}(z, n) = \{\xi_s(x) = 2 \text{ for all } s \in [nT, (n+1)T] \text{ and all } x \in Nz + (-N, N]^d \cap \mathbb{Z}^d\}$$

occurs. The same arguments as in the proofs of Lemmas 5–8 imply that

$$\mathbb{P}((\pm e_i, 1) \text{ is not good for some } i = 1, 2, \ldots, d \mid (0, 0) \text{ is good}) \leq C_4 N^d \exp(-\sqrt{N})$$

for some suitable $C_4 < \infty$; thus, there again exists a coupling such that the set of good sites contains the set of wet sites of an oriented site percolation process with parameter $1 - \gamma$ provided the parameter N is large. The reason for introducing this new definition of good sites is that the presence of a type-1 individual at vertex $x \in Nz + (-N, N]^d$ at time nT now implies the existence of a generalized dry path from some $(z_0, 0)$ to (z, n). In particular, the problem of extinction reduces to proving the lack of percolation of the generalized dry sites when $\gamma > 0$ is small enough. To do this, the first step is to prove the existence of $C_5 < \infty$ and $\gamma_5 > 0$ such that

$$\mathbb{P}(D_0(\mathcal{G}_2) \not\subset ([-m, m]^d \times \mathbb{R}) \cap \mathcal{H}) \leq C_5 \exp(-\gamma_5 m) \quad \text{for small } \gamma > 0. \quad (18)$$

This is essentially Theorem 5 of Durrett [6] which states the result for the regular dry cluster, but his proof easily extends to our case. To begin with, we define the collection of cubes

$$Q = \left\{ r \in \mathbb{R}^{d+1} : \sum_{i=1,2,\ldots,d+1} |r_i| \leq 1 \right\} \quad \text{and} \quad Q(z, n) = (z, n) + Q \quad \text{for } (z, n) \in \mathcal{H},$$

and orient the faces of the cubes by assigning the value $+1$ to the $d + 1$ faces at the top of each cube, and the value -1 to the $d + 1$ faces at the bottom. Then we define the contour of the regular and generalized dry clusters as the algebraic sum of all the faces of the cubes that make up

$$\mathcal{D}_0(\mathcal{G}_i) = \cup\{Q(z, n) : (z, n) \in D_0(\mathcal{G}_i)\} \quad \text{for } i = 1, 2. \quad (19)$$

To prove (18), the first ingredient is to observe that each contour has an equal number of plus and minus faces and sites below the plus faces must be closed. This holds regardless of the geometry of the underlying oriented graph so Lemma 4 of [6] directly applies: if the contour has length m then there is a set of at least $m(2(d+1))^{-1}$ sites that must be closed. The second ingredient is given by Lemmas 5–6 of [6] which state that there exist $C_6 < \infty$ and $\mu > 0$ such that the number of contours with m faces is bounded by $C_6 \mu^m$. The result is proved for the regular dry cluster, but again it is straightforward to extend the proof to the generalized dry cluster. To see this, the key is simply to observe that any cube that makes up the dry region (19) associated with the generalized dry cluster has at least one corner in common with another such cube. Then, following the lines of the proof of Theorem 5 of [6] and using the previous two estimates, we obtain

$$\mathbb{P}(D_0(\mathcal{G}_2) \not\subset ([-m, m]^d \times \mathbb{R}) \cap \mathcal{H}) \leq \sum_{k=m}^{\infty} C_6 \mu^k \gamma^{k(2(d+1))^{-1}},$$

from which (18) follows for $\gamma < \mu^{-2(d+1)}$. To deduce from (18) the lack of percolation of the dry sites, we denote by W_n and $\bar{W}_n$ the sets of wet sites at level n in the coupled percolation processes having the same sets of open sites except at level 0 where

$$W_0 = \{0\} \quad \text{and} \quad \bar{W}_0 = \{z \in \mathbb{Z}^d : z_1 + z_2 + \cdots + z_d \text{ is even}\}.$$

The proof of Lemma 8 of [6] gives the existence of a constant $a > 0$ such that

$$\bar{W}_n \cap B_2(0, an) \subset W_n \quad \text{for all large } n \text{ on the event } \{W_n \neq \varnothing \text{ for all } n \geq 0\}. \quad (20)$$

Following the proof of Lemma 11 of [6], and using (18) and (20), we obtain

$$\lim_{m \to \infty} \mathbb{P}\big(\text{a generalized dry path starting at some } (z, 0)$$
$$\text{intersects } \big(B_2\big(0, \tfrac{1}{2}na\big) \times \{n\}\big) \cap \mathcal{H} \text{ for some } n \geq m\big) = 0$$

on the event $\{W_n \neq \varnothing \text{ for all } n \geq 0\}$. Through the coupling with process (10), we deduce the existence of an in-all-directions expanding region containing only good sites, and, thus, the existence of an in-all-directions expanding region which is void of type 1.

From Lemma 9, we deduce that type 2 wins for process (10); therefore, the same holds for process (2) according to Lemma 4, which completes the proof of Theorem 2.

4. Selfish-selfish interactions

This section is devoted to the proof of Theorem 3 which claims the existence of cases of selfish-selfish interactions for which the most selfish type outcompetes the other type whenever starting with infinitely many representatives. The key to our proof is to first show that type 2 wins for the process with parameters $a_1 < 1$ and $a_2 = 1$, which we denote by $(\bar{\eta}_t)$, and then use a perturbation argument to extend the result to a neighborhood of this parameter point. First, we note that, under the assumption that $a_2 = 1$, the dynamics of process (2) reduce to

$$L_{\bar{\eta}} f(\bar{\eta}) = \sum_{x \in \mathbb{Z}^d} \mathbf{1}\{f_2(x, \bar{\eta}) = 0\}[f(\bar{\eta}_{x,1}) - f(\bar{\eta})]$$
$$+ \sum_{x \in \mathbb{Z}^d} \frac{(1 - a_1) f_2(x, \bar{\eta})}{a_1 f_1(x, \bar{\eta}) + (1 - a_1) f_2(x, \bar{\eta})} [f(\bar{\eta}_{x,2}) - f(\bar{\eta})]. \quad (21)$$

Invasion of type-2 individuals in the particular case when $a_2 = 1$ is established in the following lemma through a comparison with the Richardson model introduced in [16].

Lemma 10. Let $\delta > 0$. Then, for all sufficiently large N,

$$\mathbb{P}(\bar{\eta}_{C_7 N}(x) = 1 \text{ for some } x \in (-2N, 2N]^d \mid \bar{\eta}_0(x) = 2 \text{ for all } x \in (-N, N]^d) \leq \delta$$

for some constant $C_7 < \infty$ that does not depend on δ.

Proof. First, we observe that, given that vertex x has at least one neighbor of type 2 at the time it is updated, it flips from type 1 to type 2 with probability

$$\frac{(1 - a_1) f_2(x, \bar{\eta})}{a_1 f_1(x, \bar{\eta}) + (1 - a_1) f_2(x, \bar{\eta})} \geq \varepsilon := \frac{1 - a_1}{a_1(2d - 1) + (1 - a_1)} > 0,$$

while it flips from type 2 to type 1 with probability 0. This implies that, for every connected component $A \subset \mathbb{Z}^d$ containing at least two vertices, if all vertices in A are of type 2 at some time then this remains true at any later time. In particular, the set of type-2 individuals dominates stochastically the set of occupied sites in the Richardson model (ξ_t), i.e. the contact process with

no death, with birth parameter $\varepsilon (2d)^{-1} > 0$ under certain initial conditions. More precisely, both processes can be constructed from the same graphical representation in such a way that

$$\mathbb{P}(\xi_t \subset \{x : \bar{\eta}_t(x) = 2\} \text{ for all } t \geq 0 \mid \xi_0 \subset \{x : \bar{\eta}_0(x) = 2\}) = 1$$

whenever the initial set of type-2 individuals is connected and has at least two vertices. This, together with the shape theorem for the Richardson model [16], implies that there exists a finite constant $C_7 < \infty$ that depends only on the parameter ε such that, for all $\delta > 0$,

$$\mathbb{P}(\bar{\eta}_{C_7 N}(x) = 1 \text{ for some } x \in (-2N, 2N]^d \mid \bar{\eta}_0(x) = 2 \text{ for all } x \in (-N, N]^d)$$
$$\leq \mathbb{P}(\xi_{C_7 N} \not\supset (-2N, 2N]^d \mid \xi_0 = (-N, N]^d)$$
$$\leq \delta$$

for all sufficiently large N. This completes the proof.

Extinction of type-1 individuals stated in Theorem 3 follows from Lemmas 9 and 10 through the comparison of the process properly rescaled in space and time with a certain oriented site percolation process in which dry sites do not percolate. Let $T = C_7 N$, where $C_7 < \infty$ is the constant fixed in Lemma 10, and declare $(z, n) \in \mathcal{H}$ to be a good site whenever the event

$$\bar{\Omega}(z, n) = \{\bar{\eta}_s(x) = 2 \text{ for all } s \in [nT, (n+1)T] \text{ and all } x \in Nz + (-N, N]^d \cap \mathbb{Z}^d\}$$

occurs. In the particular case when $a_1 < a_2 = 1$, individuals of type 2 with at least one neighbor of type 2 cannot change; therefore, it follows from Lemma 10 that, for all $\delta > 0$,

$$\mathbb{P}((z \pm e_i, n+1) \text{ is good for all } i = 1, 2, \ldots, d \mid (z, n) \text{ is good}) \geq 1 - \delta$$

for all large N. Theorem 4.3 of [7] again gives the existence of a coupling such that the set of good sites contains the set of wet sites of a certain oriented site percolation process with finite range of dependency and parameter $1 - \delta$. To complete the proof of Theorem 3, we show extinction of individuals of type 1 when $a_2 < 1$ is close enough to 1, relying on a perturbation argument.

Lemma 11. *Let $a_1 < 1$. Then there exists $\rho > 0$ small such that*

$$\lim_{t \to \infty} \mathbb{P}(\eta_t(x) = 1) = 0 \quad \text{for all } x \in \mathbb{Z}^d \text{ whenever } a_2 > 1 - \rho.$$

Proof. First, we fix $\gamma > 0$ such that (18) in Lemma 9 holds, and set $\delta = \gamma/2$. Then we apply Lemma 10 to obtain the existence of a scaling parameter N such that

$$\mathbb{P}(\bar{\eta}_s(x) = 1 \text{ for some } (x, s) \in (-2N, 2N]^d \times [T, 2T] \mid \bar{\eta}_s(x) = 2 \text{ for}$$
$$\text{all } (x, s) \in (-N, N]^d \times [0, T])$$
$$\leq \delta. \tag{22}$$

The scaling parameter being fixed, since the transition rates are continuous with respect to the parameters, there exist a small $\rho = \rho(N) > 0$ and a coupling of (2) and (21) such that

$$\mathbb{P}(\eta_s(x) \neq \bar{\eta}_s(x) \text{ for some } (x, s) \in (-2N, 2N]^d \times [T, 2T] \mid \eta_s(x) = \bar{\eta}_s(x) = 2 \text{ for}$$
$$\text{all } (x, s) \in (-N, N]^d \times [0, T])$$
$$\leq \delta \tag{23}$$

whenever $a_2 > 1 - \rho$. By combining (22) and (23), we obtain

$$\mathbb{P}(\eta_s(x) = 1 \text{ for some } (x, s) \in (-2N, 2N]^d \times [T, 2T] \mid \eta_s(x) = 2 \text{ for}$$
$$\text{all } (x, s) \in (-N, N]^d \times [0, T])$$
$$\leq 2\delta$$
$$= \gamma$$

whenever $a_2 > 1 - \rho$. Hence, a new application of Theorem 4.3 of [7] gives the existence of a coupling such that the set of good sites contains the set of wet sites of an oriented site percolation process with parameter $1 - \gamma$ when $a_2 > 1 - \rho$. Having such a coupling, extinction of type-1 individuals directly follows as previously from the proof of Lemma 9.

5. Strong altruistic-altruistic interactions

This section is devoted to the proof of Theorem 4 which indicates that, in contrast with the one-dimensional nearest-neighbor model, coexistence occurs in any other spatial dimensions and for any other range of interactions provided cooperation is strong enough. The first step of the proof is to show that, excluding the one-dimensional nearest-neighbor case, the critical value of the threshold contact process is strictly smaller than 1, and then use a standard coupling argument to deduce the existence of a small constant $\rho > 0$ such that type-1 individuals survive for the process with parameters $a_1 = 0$ and $a_2 = \rho$. The coexistence result will then follow from symmetry and the monotonicity of the survival probabilities with respect to the parameters. Recall that the threshold contact process with parameter α is the spin system (ξ_t) in which each occupied vertex becomes empty at rate 1, and each empty vertex with at least one occupied neighbor becomes occupied at rate α. Formally, the dynamics are described by the Markov generator

$$L_\xi f(\xi) = \sum_{x \in \mathbb{Z}^d} \alpha \, \mathbf{1}\{f_1(x, \xi) \neq 0\}[f(\xi_{x,1}) - f(\xi)] + \sum_{x \in \mathbb{Z}^d} [f(\xi_{x,0}) - f(\xi)],$$

where empty and occupied vertices are identified to states 0 and 1, respectively. The reason for looking at the threshold contact process is that, when the birth rate $\alpha = 1$, it is closely related to process (2) with parameters $a_1 = a_2 = 0$.

Lemma 12. *Assume that* $(M, d) \neq (1, 1)$. *Then*

$$\alpha_c = \inf\{\alpha \colon \mathbb{P}(\xi_t \neq \varnothing \text{ for all } t \geq 0 \mid \xi_0 = \{0\}) > 0\} < 1.$$

Proof. The rate at which a vertex becomes empty is constant while the rate at which a vertex becomes occupied is a nondecreasing function of the configuration, where the partial order on the set of configurations is the inclusion. This indicates that the threshold contact process is an attractive spin system. In particular, Theorem 2.4 of [2] implies that

$$\{\alpha \colon \mathbb{P}(\xi_t \neq \varnothing \text{ for all } t \geq 0 \mid \xi_0 = \{0\}) > 0\} \text{ is an open interval.} \tag{24}$$

In addition, Lemma 2.1 of [9], which relies on a result of [14], implies that

$$\mathbb{P}(\xi_t \neq \varnothing \text{ for all } t \geq 0 \mid \xi_0 = \{0\}) > 0 \quad \text{for } (M, d) \neq (1, 1) \text{ and } \alpha \geq 1. \tag{25}$$

Combining (24) and (25) gives $1 \in (\alpha_c, \infty)$; therefore, $\alpha_c < 1$.

Relying on Lemma 12, we now prove the existence of $\rho > 0$ small such that type-1 individuals survive for the process with parameters $a_1 = 0$ and $a_2 = \rho$, which we denote by $(\bar{\eta}_t)$. Note that, for this choice of parameters, the dynamics reduce to

$$L_{\bar{\eta}} f(\bar{\eta}) = \sum_{x \in \mathbb{Z}^d} \frac{(1-\rho) f_1(x, \bar{\eta})}{(1-\rho) f_1(x, \bar{\eta}) + \rho f_2(x, \bar{\eta})} [f(\bar{\eta}_{x,1}) - f(\bar{\eta})]$$
$$+ \sum_{x \in \mathbb{Z}^d} \mathbf{1}\{f_2(x, \bar{\eta}) \neq 0\} [f(\bar{\eta}_{x,2}) - f(\bar{\eta})]. \qquad (26)$$

Lemma 13. *For all* $(M, d) \neq (1, 1)$, *there exists* $\rho > 0$ *such that*

$$\liminf_{t \to \infty} \mathbb{P}(\bar{\eta}_t(x) = 1) > 0 \quad \text{for all } x \in \mathbb{Z}^d.$$

Proof. Thinking of vertices of type 1 as occupied by a particle and those of type 2 as empty, (26) indicates that particles die at rate at most 1, while empty vertices with at least one occupied neighbor become occupied at rate at least

$$\alpha(\rho) = \frac{1-\rho}{1 - \rho + \rho((2M+1)^d - 1)} = \frac{1-\rho}{1 + ((2M+1)^d - 2)\rho} > \alpha_c,$$

provided $\rho > 0$ is small enough since $\alpha_c < 1$ according to Lemma 12. Therefore, for $\rho > 0$ small such that the previous inequality holds, the set of type-1 vertices dominates stochastically the set of occupied sites of a supercritical threshold contact process. Hence, it suffices to prove the result for the latter: starting with infinitely many occupied vertices,

$$\liminf_{t \to \infty} \mathbb{P}(\xi_t(x) = 1) > 0 \quad \text{for all } x \in \mathbb{Z}^d \text{ and all } \alpha > \alpha_c.$$

This is a direct consequence of the complete convergence theorem for the threshold contact process established in [5] which implies that, starting with infinitely many occupied vertices, the process converges in distribution to the limit obtained when starting with all vertices occupied, which is a nontrivial translation invariant measure when $\alpha > \alpha_c$.

To deduce Theorem 4 from Lemma 13, we first invoke the monotonicity of the survival probabilities with respect to the parameters to deduce that, for $\rho > 0$ as in the lemma, type 1 survives when the parameters are $a_1 = a_2 = \rho$. By symmetry, individuals of type 2 survive as well, and so both types coexist for these parameters. Finally, invoking again the monotonicity of the survival probabilities, we deduce that coexistence occurs for all $a_1, a_2 < \rho$.

6. Strong altruistic-weak altruistic interactions

This section is devoted to the proof of Theorem 5 which gives the existence of a parameter region in which coexistence occurs for the nonspatial model but not the spatial model. The proof combines ideas from Sections 3 and 4. First, we study the process with $a_1 = 0$ and $a_2 = \frac{1}{2}$ using techniques which are reminiscent of duality techniques, though the process does not have a dual process at this parameter point. This is done through the new concept of breaking point that we introduce in this paper. Using breaking points, we construct selected dual paths as in Section 3, and using a block construction, we deduce that type 2 wins when $a_1 = 0$ and $a_2 = \frac{1}{2}$. To complete the proof of the theorem, we rely as in Section 4 on a perturbation argument.

6.1. Breaking points

First, we assume that $a_1 = 0$ and $a_2 = \frac{1}{2}$. Under this assumption, the type at each vertex x is updated at rate 1 according to the following rules.

- If vertex x is of type 2 at the time of the update then the new type is chosen uniformly at random from the set of the $2d$ nearest neighbors.

- If vertex x is of type 1 at the time of the update then it switches to type 2 whenever at least one of its nearest neighbors is of type 2.

More formally, the dynamics of the process with $a_1 = 0$ and $a_2 = \frac{1}{2}$, which we denote by $(\bar{\eta}_t)$, are described by the Markov generator

$$L_{\bar{\eta}} f(\bar{\eta}) = \sum_{x \in \mathbb{Z}^d} f_1(x, \bar{\eta})[f(\bar{\eta}_{x,1}) - f(\bar{\eta})] + \sum_{x \in \mathbb{Z}^d} \mathbf{1}\{f_2(x, \bar{\eta}) \neq 0\}[f(\bar{\eta}_{x,2}) - f(\bar{\eta})]. \quad (27)$$

We construct the process (27) graphically as follows:

- for each oriented edge $e = (x, y) \in \mathbb{Z}^d \times \mathbb{Z}^d$ with $x \sim y$,

  - we let $\Lambda(x, y)$ be a Poisson point process with intensity $(2d)^{-1}$,

  - we draw at all times $t \in \Lambda(x, y)$ a solid arrow $y \to x$ and an additional $2d - 1$ dashed arrows from the other neighbors of x to vertex x to indicate that if x is of type 2 then it mimics vertex y, whereas if it is of type 1 then it switches to type 2 whenever at least one of its neighbors is of type 2.

Definition 2. We call (x, t) where $t \in \Lambda(x, y)$ a breaking point if

$$\sup\{\Lambda(y, x) \cap (0, t)\} = \sup\{(\Lambda(x, z) \cup \Lambda(y, z)) \cap (0, t) : z \sim x \text{ or } z \sim y\}$$

and the set $\Lambda(y, x) \cap (0, t)$ is nonempty.

In words, (x, t), $t \in \Lambda(x, y)$, is a breaking point if the next solid arrow pointing at either vertex x or vertex y we see going down the graphical representation is directed from x to y. We refer the reader to the right-hand diagram of Figure 2 for an illustration of the graphical representation that includes the breaking points, which we represent with filled circles. The reason for looking at breaking points is that they allow us, as in the proof of Theorem 2, to construct selected dual paths with a drift towards regions void of type-1 individuals and, thus, to prove invasion of type 2. This good property of the breaking points is expressed in the following lemma.

Lemma 14. *Assume that (x, t) is a breaking point. Then*

$$\bar{\eta}_t(x) = 1 \quad \text{if and only if} \quad \bar{\eta}_t(z) = 1 \quad \text{for all } z \sim x.$$

Proof. Assume that $t \in \Lambda(x, y)$, and let $s = \sup\{\Lambda(y, x) \cap (0, t)\}$. Then

$$\bar{\eta}_s(x) = 2 \quad \Longrightarrow \quad \bar{\eta}_s(y) = 2 \text{ because there is a solid arrow } x \to y \text{ at time } s \qquad (28a)$$
$$\Longrightarrow \quad \bar{\eta}_r(y) = 2 \text{ for all } r \in (s, t) \text{ since no arrow points at } \{y\} \times (s, t) \quad (28b)$$
$$\Longrightarrow \quad \bar{\eta}_t(x) = 2 \text{ because there is a solid arrow } y \to x \text{ at time } t. \qquad (28c)$$

From (28), we deduce that

$$\bar\eta_t(x) = 1 \implies \bar\eta_t(x) = 1 \text{ and } \bar\eta_s(x) = 1$$
$$\implies \bar\eta_t(x) = 1 \text{ and } \bar\eta_{t-}(x) = 1 \text{ because no arrow points at } \{x\} \times (s, t)$$
$$\implies \bar\eta_t(z) = 1 \text{ for all } z \sim x,$$

where the last implication follows from the evolution rules of the process. This proves the necessary condition. The sufficient condition is trivial and we point out that it is always true provided an arrow points at (x, t), whereas the reverse holds because (x, t) is a breaking point.

Motivated by Lemma 14, we now construct from the graphical representation a certain system of branching coalescing random walks that we will call a dual process as it allows us to obtain information about the type of a given vertex, say x, at the current time, say T, based on the configuration at earlier times. To define the dual process starting at (x, T), the first step is to remove from the graphical representation all the dashed arrows that point at a space–time point which is not a breaking point. Then we say that there is a dual path from (x, T) down to $(y, T - s)$, which we again write as $(x, T) \downarrow (y, T - s)$, whenever there are sequences of times and vertices

$$s_0 = T - s < s_1 < \cdots < s_{m+1} = T \quad \text{and} \quad x_0 = y, x_1, \ldots, x_m = x$$

such that the following two conditions hold:

1. for $i = 1, 2, \ldots, m$, there is an arrow from x_{i-1} to x_i at time s_i, and

2. for $i = 0, 1, \ldots, m$, there is no arrow that points at the segments $\{x_i\} \times (s_i, s_{i+1})$.

We point out that condition 1 above must hold after removal of the dashed arrows that are not directed to a breaking point. The dual process starting at (x, T) is the set-valued process

$$\hat\eta_s(x, T) = \{y \in \mathbb{Z}^d : (x, T) \downarrow (y, T - s)\} \quad \text{for all } 0 \le s \le T.$$

See the right-hand diagram of Figure 2 where the dual process is represented by thick lines. One of the keys to proving Theorem 5 is the following duality relationship which is simply obtained by applying Lemma 14 at each of the breaking points that occurs along the dual process:

$$\bar\eta_T(x) = 1 \quad \text{only if} \quad \bar\eta_{T-s}(y) = 1 \quad \text{for all } y \in \hat\eta_s(x, T). \tag{29}$$

6.2. Selected dual paths

To bound the probability that space–time point (x, T) is of type 1, we follow the same approach as in Section 3 and construct a process $(Y_t(x))$ that keeps track of a specific particle in the dual process. Recall that

$$H_i := \{z = (z_1, z_2, \ldots, z_d) \in \mathbb{R}^d : z_i = 0\} \quad \text{for } i = 1, 2, \ldots, d.$$

In addition, we let $T_i := i\, c_2 N$ for $i = 0, 1, \ldots, d$, where

$$c_2 := 4\varepsilon^{-1} \quad \text{with} \quad \varepsilon := 2d(1 - e^{-1/2})^2 e^{-(2d-1)} > 0$$

and where N is a scaling parameter that will be fixed later. The choice of ε will become clear in the proof of Lemma 15 below. The dual path starts at $Y_0(x) = x$. To construct this path at later times, assuming that it is defined up to time t, we introduce

$$s(t) = \inf\{s > t : T - s \in \Lambda(Y_t(x), y) \text{ for some } y \sim Y_t(x)\}. \tag{30}$$

For all $s \in (t, s(t))$, we set $Y_s(x) = Y_t(x)$, while at time $s(t)$ we have the following alternatives.

- If $(Y_t(x), T - s(t))$ is a breaking point and $t \in (T_{i-1}, T_i)$ for some $i = 1, 2, \ldots, d$, then

$$Y_{s(t)}(x) \sim Y_t(x) \quad \text{and} \quad \text{dist}(Y_{s(t)}(x), H_i) < \text{dist}(Y_t(x), H_i).$$

- Otherwise, $Y_{s(t)}(x) = y$ is the neighbor for which the infimum in (30) is reached.

In words, the process goes backwards in time down the graphical representation following the solid arrows from tip to tail except when a breaking point is encountered, in which case the process crosses the arrow, either dashed or solid, that makes it closer to the hyperplane H_i between the times T_{i-1} and T_i. After time T_d only the solid arrows are crossed. The process is thus embedded in the dual process which, together with relationship (29), implies that

$$\bar{\eta}_T(x) = 1 \quad \text{only if} \quad Y_T(x) = 1. \tag{31}$$

Motivated by condition (31), the next step is to prove the analog of Lemma 7 for this new selected dual path. More precisely, we have the following lemma.

Lemma 15. *Let $x \in (-2N, 2N]^d$. Then there exists $C_8 < \infty$ such that*

$$\mathbb{P}(Y_t(x) \notin (-4N, 4N]^d \text{ for some } t < C_8 N \text{ or } Y_{C_8 N}(x) \notin (-N, N]^d) \leq C_9 \exp(-\gamma_9 \sqrt{N})$$

for suitable $C_9 < \infty$ and $\gamma_9 > 0$, and all sufficiently large N.

Proof. To begin with, we call $s \in (0, T - 1)$ a breaking time whenever

1. there is a unique neighbor y of vertex $Y_s(x)$ such that there is a solid arrow $y \to Y_s(x)$ in the time interval $(T - s, T - s - \frac{1}{2})$,

2. there is a unique solid arrow $Y_s(x) \to y$ in the time interval $(T - s - \frac{1}{2}, T - s - 1)$,

3. no other solid arrows point at $Y_s(x)$ or y in the time interval $(T - s, T - s - 1)$.

Note that these three conditions imply that the dual path crosses exactly one breaking point between dual times s and $s + 1$. Moreover, a simple calculation shows that

$$\mathbb{P}(s \text{ is a breaking time}) = 2d(1 - e^{-1/2})^2 e^{-(2d-1)} = \varepsilon > 0.$$

Since, in addition, disjoint parts of the graphical representations are independent, it follows that the discrete-time process $(Y_n(x): n \in (0, T) \cap \mathbb{Z})$ is a Markov chain that satisfies

$$\mathbb{E}(\text{dist}(Y_{n+1}(x), H_i) - \text{dist}(Y_n(x), H_i))$$
$$\leq \varepsilon \mathbb{E}(\text{dist}(Y_{n+1}(x), H_i) - \text{dist}(Y_n(x), H_i) \mid n \text{ is a breaking time})$$
$$\leq -\varepsilon$$

for all $n \in (T_{i-1}, T_i) \cap \mathbb{Z}$. This is the analog of (14). In particular, the rest of the proof follows from the exact same arguments as in the proofs of Lemmas 5–7.

6.3. Block construction

To complete the proof of Theorem 5, we let $T = C_8 N$ and declare site $(z, n) \in \mathcal{H}$ to be a good site whenever the following event occurs:

$$\bar{\Omega}(z, n) = \{\bar{\eta}_s(x) = 2 \text{ for all } s \in [nT, (n+1)T] \text{ and all } x \in Nz + (-N, N]^d \cap \mathbb{Z}^d\}.$$

Lemma 15 and (31) imply that, for all $\delta > 0$,

$$\mathbb{P}((z \pm e_i, n+1) \text{ is good for all } i = 1, 2, \ldots, d \mid (z, n) \text{ is good}) \geq 1 - \delta,$$

provided the scale parameter N is sufficiently large. This, together with the exact same arguments as in the proofs of Lemmas 8 and 9, implies that type 2 wins for process (27). Finally, using the continuity of the transition rates with respect to the parameters and the same approach as in the proof of Lemma 11, we deduce the existence of $\rho > 0$ small such that type 2 wins as well for process (2) when $a_1 < \rho$ and $a_2 > \frac{1}{2} - \rho$. This completes the proof.

Acknowledgement

References

[1] ATHREYA, S. R. AND SWART, J. M. (2005). Branching-coalescing particle systems. *Prob. Theory Relat. Fields* **131,** 376–414.

[2] BEZUIDENHOUT, C. AND GRAY, L. (1994). Critical attractive spin systems. *Ann. Prob.* **22,** 1160–1194.

[3] BRAMSON, M. AND DURRETT, R. (1988). A simple proof of the stability criterion of Gray and Griffeath. *Prob. Theory Relat. Fields* **80,** 293–298.

[4] BRAMSON, M. AND GRIFFEATH, D. (1981). On the Williams-Bjerknes tumour growth model. I. *Ann. Prob.* **9,** 173–185.

[5] COX, J. T. AND DURRETT, R. (1991). Nonlinear voter models. In *Random Walks, Brownian Motion, and Interacting Particle Systems* (Progress Prob. **28**), Birkhäuser, Boston, MA, pp. 189–201.

[6] DURRETT, R. (1992). Multicolor particle systems with large threshold and range. *J. Theoret. Prob.* **5,** 127–152.

[7] DURRETT, R. (1995). Ten lectures on particle systems. In *Lectures on Probability Theory* (Saint-Flour, 1993; Lecture Notes Math. **1608**), Springer, Berlin, pp. 97–201.

[8] DURRETT, R. AND LEVIN, S. (1994). The importance of being discrete (and spatial). *Theoret. Pop. Biol.* **46,** 363–394.

[9] HANDJANI, S. J. (1999). The complete convergence theorem for coexistent threshold voter models. *Ann. Prob.* **27,** 226–245.

[10] KANG, Y. AND LANCHIER, N. (2012). The role of space in the exploitation of resources. *Bull. Math. Biol.* **74,** 1–44.

[11] KRONE, S. M. AND NEUHAUSER, C. (1997). Ancestral processes with selection. *Theoret. Pop. Biol.* **51,** 210–237.

[12] LANCHIER, N. AND NEUHAUSER, C. (2009). Spatially explicit non-Mendelian diploid model. *Ann. Appl. Prob.* **19,** 1880-1920.

[13] LIGGETT, T. M. (1985). *Interacting Particle Systems* (Fundamental Principles Math. Sci. **276**). Springer, New York.

[14] LIGGETT, T. M. (1994). Coexistence in threshold voter models. *Ann. Prob.* **22,** 764–802.

[15] NEUHAUSER, C. AND PACALA, S. W. (1999). An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann. Appl. Prob.* **9,** 1226–1259.

[16] RICHARDSON, D. (1973). Random growth in a tessellation. *Proc. Camb. Phil. Soc.* **74,** 515–528.