The Annals of Applied Probability
2013, Vol. 23, No. 6, 2538–2559
DOI: 10.1214/12-AAP910
© Institute of Mathematical Statistics, 2013

FIXATION IN THE ONE-DIMENSIONAL AXELROD MODEL

BY NICOLAS LANCHIER[1] AND STYLIANOS SCARLATOS[2]

Arizona State University and University of Patras

The Axelrod model is a spatial stochastic model for the dynamics of cultures which includes two important social factors: social influence, the tendency of individuals to become more similar when they interact, and homophily, the tendency of individuals to interact more frequently with individuals who are more similar. Each vertex of the interaction network is characterized by its culture, a vector of F cultural features that can each assumes q different states. Pairs of neighbors interact at a rate proportional to the number of cultural features they have in common, which results in the interacting pair having one more cultural feature in common. In this article, we continue the analysis of the Axelrod model initiated by the first author by proving that the one-dimensional system fixates when $F \leq cq$ where the slope satisfies the equation $e^{-c} = c$. In addition, we show that the two-feature model with at least three states fixates. This last result is sharp since it is known from previous works that the one-dimensional two-feature two-state Axelrod model clusters.

1. Introduction. The Axelrod model is one of the most popular agent-based models of cultural dynamics. In addition to a spatial structure, which is modeled through a graph in which vertices represent individuals and edges potential dyadic interactions between two individuals, it includes two important social factors: social influence and homophily. The former is the tendency of individuals to become more similar when they interact, while the latter is the tendency of individuals to interact more frequently with individuals who are more similar. Note that the voter model [5, 7] accounts for social influence since an interaction between two individuals results in a perfect agreement between them. The voter model, however, excludes homophily. To also account for this factor, one needs to be able to define a certain opinion or cultural distance between any two individuals through which the frequency of the interactions between the two individuals can be measured. In the model proposed by political scientist Robert Axelrod [1], each individual is characterized by her opinions on F different cultural features, each of which assumes q possible states. Homophily is modeled by assuming that pairs of neighbors interact at a rate equal to the fraction of cultural features for which they agree,

---

Received March 2012; revised November 2012.

[1]Supported in part by NSF Grant DMS-10-05282.

[2]Supported in part by a grant from the GSRT, Greek Ministry of Development, for the project "Complex Matter", awarded under the auspices of the ERA Complexity Network.

MSC2010 subject classifications. 60K35.

Key words and phrases. Interacting particle systems, Axelrod model, random walks, fixation.

and social influence by assuming that, as a result of the interaction, one of the cultural features for which members of the interacting pair disagree (if any) is chosen uniformly at random, and the state of one of both individuals is set equal to the state of the other individual for this cultural feature. More formally, the Axelrod model on the one-dimensional lattice is the continuous-time Markov chain whose state space consists of all spatial configurations

$$\eta : \mathbb{Z} \longrightarrow \{1, 2, \dots, q\}^F$$

that map the vertex set viewed as the set of all individuals into the set of cultures. To describe the dynamics of the Axelrod model, it is convenient to introduce

$$F(x, y) := \frac{1}{F} \sum_{i=1}^{F} \mathbf{1}\{\eta(x, i) = \eta(y, i)\},$$

where $\eta(x, i)$ refers to the ith coordinate of the vector $\eta(x)$, which denotes the fraction of cultural features the two vertices x and y share. To describe the elementary transitions of the spatial configuration, we also introduce the operator $\sigma_{x,y,i}$ defined on the set of configurations by

$$(\sigma_{x,y,i}\eta)(z, j) := \begin{cases} \eta(y, i), & \text{if } z = x \text{ and } j = i, \\ \eta(z, j), & \text{otherwise,} \end{cases} \qquad \text{for } x, y \in \mathbb{Z} \text{ and } i \in \{1, 2, \dots, F\}.$$

In other words, configuration $\sigma_{x,y,i}\eta$ is obtained from configuration η by setting the ith feature of the individual at vertex x equal to the ith feature of the individual at vertex y and leaving the state of all the other features in the system unchanged. The dynamics of the Axelrod model is then described by the Markov generator L defined on the set of cylinder functions by

$$Lf(\eta) := \sum_{|x-y|=1} \sum_{i=1}^{F} \frac{1}{2F} \left[\frac{F(x, y)}{1 - F(x, y)} \right] \mathbf{1}\{\eta(x, i) \neq \eta(y, i)\} [f(\sigma_{x,y,i}\eta) - f(\eta)].$$

The expression of the Markov generator indicates that the conditional rate at which the ith feature of vertex x is set equal to the ith feature of vertex y given that these two vertices are nearest neighbors that disagree on their ith feature can be written as

$$\frac{1}{2F} \left[\frac{F(x, y)}{1 - F(x, y)} \right] = F(x, y) \times \frac{1}{F(1 - F(x, y))} \times \frac{1}{2},$$

which, as required, is equal to the fraction of features both vertices have in common, which is the rate at which the vertices interact, times the reciprocal of the number of features for which both vertices disagree, which is the probability that any of these features is the one chosen for update, times the probability one half that vertex x rather than vertex y is chosen to be updated. Note that, when the number of features $F = 1$, the system is static, while when the number of states per feature $q = 1$ there is only one possible culture. Also, to avoid trivialities, we assume from now on that the two parameters of the system are strictly larger than one.

The main question about the Axelrod model is whether the system fluctuates and evolves to a global consensus or gets trapped in a highly fragmented configuration. To define this dichotomy rigorously, we say that the system fluctuates whenever

$$P\big(\eta_t(x,i) \text{ changes value at arbitrary large } t\big) = 1 \qquad \text{for all } x \in \mathbb{Z} \text{ and } i \in \{1,2,\ldots,F\} \tag{1}$$

and fixates if there exists a configuration η_∞ such that

$$P\big(\eta_t(x,i) = \eta_\infty(x,i) \text{ eventually in } t\big) = 1 \qquad \text{for all } x \in \mathbb{Z} \text{ and } i \in \{1,2,\ldots,F\}. \tag{2}$$

In other words, fixation means that the culture of each individual is only updated a finite number of times, so fluctuation (1) and fixation (2) exclude each other. We define convergence to a global consensus mathematically as a clustering of the system, that is,

$$\lim_{t\to\infty} P\big(\eta_t(x,i) = \eta_t(y,i)\big) = 1 \qquad \text{for all } x, y \in \mathbb{Z} \text{ and } i \in \{1,2,\ldots,F\}. \tag{3}$$

Note that whether the system fluctuates or fixates depends not only on the number of cultural features and the number of states per feature, but also on the initial distribution. Indeed, regardless of the parameters, the system starting from a configuration in which all the individuals agree for a given cultural feature while the states at the other cultural features are independent and occur with the same probability always fluctuates. On the other hand, regardless of the parameters, the system starting from a configuration in which all the even sites share the same culture and all the odd sites share another culture which is incompatible with the one at even sites always fixates. Also, we say that fluctuation/fixation occurs for a given pair of parameters if the one-dimensional system with these parameters fluctuates/fixates when starting from the distribution π_0 in which the states of the cultural features within each vertex and among different vertices are independent and uniformly distributed. We also point out that neither fluctuation implies clustering nor fixation excludes clustering in general. Indeed, the voter model in dimensions larger than or equal to three for which coexistence occurs is an example of spin system that fluctuates but does not cluster while the biased voter model [2, 3] is an example of spin system that fixates and clusters. In spite of these counter-examples, we conjecture that fluctuation implies clustering and fixation excludes clustering for the one-dimensional Axelrod model starting from the distribution π_0.

We now give a brief review of the previous results about the one-dimensional Axelrod model and state the new results proved in this article. Since two neighbors are more likely to interact as the number of cultural features increases and the number of states per feature decreases, one expects the phase transition between the fluctuation/clustering regime and the fixation/no clustering regime to be an increasing function in the F-q plane. The numerical simulations together with the mean-field approximation of [11] suggest that the system starting from π_0:

- exhibits consensus (clustering) when $q < F$ and
- gets trapped in a highly fragmented configuration (no clustering) when $F < q$.

Looking now at analytical results, the first result in [8] states that the one-dimensional, two-feature, two-state Axelrod model clusters. The second result deals with the system on a large but finite interval, and indicates that, for a certain subset of the parameter region, the system gets trapped in a random configuration in which the expected number of cultural domains scales like the number of vertices. This strongly suggests fixation of the infinite system in this parameter region, which we prove in this paper. Shortly after, Lanchier and Schweinsberg [9] realized that the analysis of the Axelrod model can be greatly simplified using a coupling to translate problems about the model into problems about a certain system of random walks. To visualize this coupling, think of each spatial configuration as a q-coloring of the set $\mathbb{Z} \times \{1, 2, \dots, F\}$ and

$$\text{put a particle at } (u, i) \text{ whenever } \eta(u - 1/2, i) \neq \eta(u + 1/2, i) \tag{4}$$

for all $u \in \mathbb{Z} + 1/2$ and all cultural features i. We call u a blockade when it contains F particles, or equivalently when the two individuals on each side of u completely disagree. When the number of states per feature $q = 2$, Lanchier and Schweinsberg [9] proved that construction (4) induces a system of annihilating symmetric random walks that has a certain site recurrence property, which is equivalent to fluctuation of the Axelrod model, when starting from π_0. From this property, they also deduced extinction of the blockades and clustering, thus extending the first result of [8] to the model with two states per feature and any number of features. In contrast, the present paper deals with the fixation part of the conjecture and extends the second result of [8] by again using the random walk representation induced by (4). The first step is to prove that, for all values of the parameters, construction (4) induces a system of random walks in which collisions result independently in either annihilation or coalescence with some specific probabilities. Coalescing events only occur when the number of states $q > 2$. This is then combined with large deviation estimates for the initial distribution of particles to obtain survival of the blockades when starting from π_0 in the parameter region described in the second result of [8]. This not only implies fixation of the infinite system, but also excludes clustering so the system gets trapped in a highly fragmented configuration.

THEOREM 1. *Assume that*

$$\omega(q, F) := q\left(1 - \frac{1}{q}\right)^F - F\left(1 - \frac{1}{q}\right) > 0. \tag{5}$$

Then, fixation (2) *occurs and clustering* (3) *does not occur.*

Interestingly, though the second result in [8] relies on a coupling between the Axelrod model and a certain urn problem along with some combinatorial techniques that strongly differ from the techniques in our proof, both approaches lead to the same sufficient condition (5). The set of parameters described implicitly in condition (5) corresponds to the triangular set of crosses in the two diagrams of Figure 1, which we obtained using a computer program. The picture suggests that this parameter region is (almost) equal to the set of parameters below a certain straight line going through the origin. To find the asymptotic slope, observe that if $F = cq$, then

$$\lim_{q\to\infty} q^{-1}\omega(q, F) = \lim_{q\to\infty}\left(1 - \frac{1}{q}\right)^{cq} - c\left(1 - \frac{1}{q}\right) = e^{-c} - c.$$

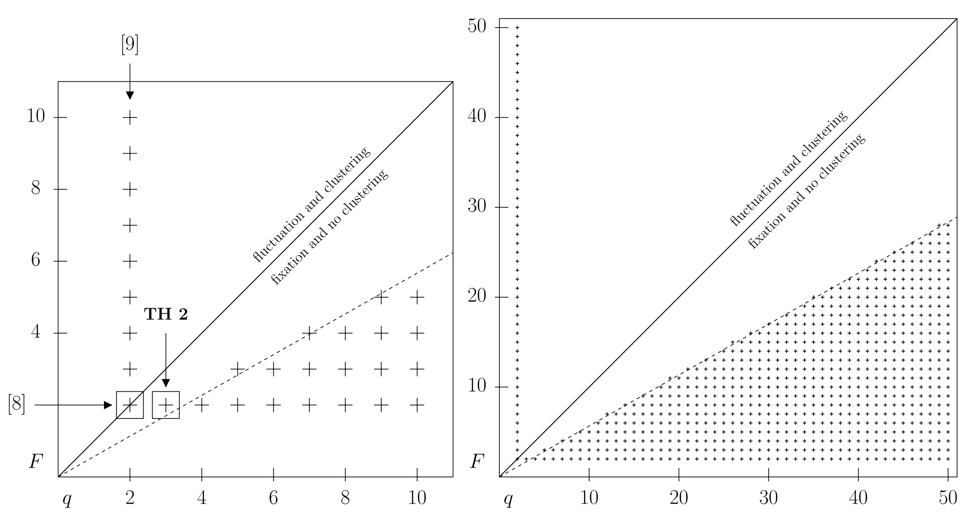


FIG. 1. *Phase diagram of the one-dimensional Axelrod model in the F-q plane. The diagram on the left-hand side is simply an enlargement of the diagram on the right-hand side that focuses on small parameters. The continuous straight line with equation $F = q$ is the transition curve conjectured in* [11]. *The set of crosses is the set of parameters for which the conjecture has been proved analytically: the vertical line of crosses on the left-hand side of the diagrams is the set of parameters for which fluctuation and clustering have been proved in* [9] *while the triangular set of crosses is the set of parameters such that $\omega(q, F) > 0$ for which fixation is proved in Theorem* 1. *The dashed line is the straight line with equation $F = cq$ where the slope c is such that $c = e^{-c}$.*

In other respects, if $e^{-c} = c$, then we have

$$\begin{aligned}(cq-1)\ln\left(1-\frac{1}{q}\right)-\ln(c) &= (1-cq)\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{1}{q}\right)^n + c \\ &= \sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{1}{q}\right)^n - \sum_{n=0}^{\infty}\frac{c}{n+1}\left(\frac{1}{q}\right)^n + c \\ &= \sum_{n=1}^{\infty}\left(\frac{1}{n}-\frac{c}{n+1}\right)\left(\frac{1}{q}\right)^n > 0\end{aligned}$$

from which we deduce that

$$cq\ln\left(1-\frac{1}{q}\right) > \ln(c)+\ln\left(1-\frac{1}{q}\right) \quad \text{and} \quad \left(1-\frac{1}{q}\right)^{cq} > c\left(1-\frac{1}{q}\right).$$

This proves that the condition in the theorem holds for $F = cq$ and so all $F \le cq$ since ω is decreasing with respect to its second variable. In particular, fixation occurs whenever

$$F \le cq \qquad \text{where } c \approx 0.567 \text{ satisfies } e^{-c} = c.$$

See Figure 1 for a picture of the straight line with equation $F = cq$. Finally, though $\omega(3, 2) = 0$ and therefore Theorem 1 does not imply fixation for the two-feature three-state Axelrod model, our approach can be improved to also obtain fixation in this case.

THEOREM 2. *The conclusion of Theorem 1 holds whenever $F = 2$ and $q = 3$.*

Note that this fixation result is sharp since the first result in [8] gives fluctuation and clustering of the two-feature two-state Axelrod model in one dimension. In particular, the two-feature model fixates if and only if the number of states per feature $q > 2$. To conclude, we note that, in contrast with the techniques introduced in [9] that heavily relies on the fact that the system starts from π_0, our proof of Theorem 1 easily extends to show that, starting from more general product measures, the one-dimensional system fixates under a certain assumption stronger than (5). However, the estimates of Lemmas 3 and 6, and consequently the condition for fixation, in this more general context become very messy while the proof does not bring any new interesting argument. Therefore, we focus for simplicity on the most natural initial distribution π_0.

2. Coupling with annihilating-coalescing random walks. As pointed out in [8], one key to understanding the Axelrod model is to keep track of the disagreements between neighbors rather than the actual set of opinions of each individual. When the number of states per feature $q = 2$, this results in a collection of nonindependent systems of annihilating symmetric random walks. Lanchier and Schweinsberg [9] have recently studied these systems of random walks in detail and deduced from their analysis that the two-state Axelrod model clusters in one dimension. When the number of states per feature is larger than two, these systems are more complicated because each collision between two random walks can result in either both random walks annihilating or both random walks coalescing. In this section, we recall the connection between the Axelrod model and systems of symmetric random walks, and complete the construction given in [9] to also include the case $q > 2$ in which coalescing events take place.

To begin with, we think of each edge of the graph as having F levels, and place a particle on an edge at level i if and only if the two individuals that this edge connects disagree on their ith feature. More precisely, we define the process

$$\xi_t(u,i) := \mathbf{1}\{\eta_t(u-1/2,i) \neq \eta_t(u+1/2,i)\} \qquad \text{for all } u \in \mathbb{D} := \mathbb{Z} + 1/2$$

and place a particle at site $u \in \mathbb{D}$ at level i whenever $\xi_t(u,i) = 1$. To describe this system, it is convenient to also introduce the process that keeps track of the number of particles per site,

$$\zeta_t(u) := \sum_{i=1}^{F} \xi_t(u,i) \qquad \text{for all } u \in \mathbb{D},$$

and to call site u a j-site whenever it contains a total of j particles: $\zeta_t(u) = j$. To understand the dynamics of these particles, the first key is to observe that, since each interaction between two individuals is equally likely to affect the culture of any of these two individuals, each particle moves one unit to the right or one unit to the left with equal probability one half. Because the rate at which two neighbors interact is proportional to the number of cultural features they have in common, a particle at (u,i) jumps at a rate that depends on the total number of particles located at site u, which induces systems of particles which are not independent. More precisely, since two adjacent vertices that disagree on exactly j of their features, and therefore are connected by an edge that contains a pile of j particles, interact at rate $1 - j/F$, the fraction of features they share, conditional on the event that u is a j-site, each particle at site u jumps at rate

$$r(j) := \left(1 - \frac{j}{F}\right)\frac{1}{j} = \frac{1}{j} - \frac{1}{F} \qquad \text{for } j \neq 0, \tag{6}$$

which represents the rate at which both vertices interact times the probability that any of the j particles is the one selected to jump. Motivated by (6), the particles at site u are said to be active if the site has less than F particles, and frozen if the site has F particles, in which case we call u a blockade. To complete the construction of these systems of random walks, the last step is to understand the outcome of a collision between two particles. Assume that (u,i) and $(u+1,i)$ are occupied at

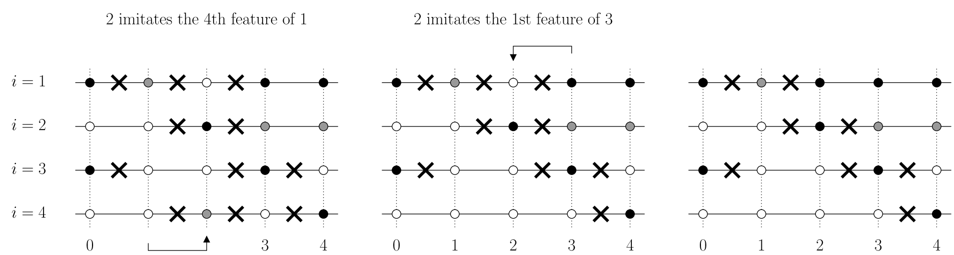


FIG. 2. *Coupling between the Axelrod model and annihilating-coalescing random walks.*

time $t-$ and that the particle at (u, i) jumps one unit to the right at time t, an event that we call a collision and that we denote by

$$(u, i) \longrightarrow (u + 1, i) \qquad \text{at time } t.$$

This happens when the individual at $x := u + 1/2$ disagrees with her two nearest neighbors on her ith feature at time $t-$ and imitates the ith feature of her left neighbor at time t. This collision results in two possible outcomes. If the individuals at x and $x + 1$ agree on their ith feature just after the update, or equivalently the individuals at $x - 1$ and $x + 1$ agree on their ith feature just before the update, then $(u + 1, i)$ becomes empty so both particles annihilate, which we write

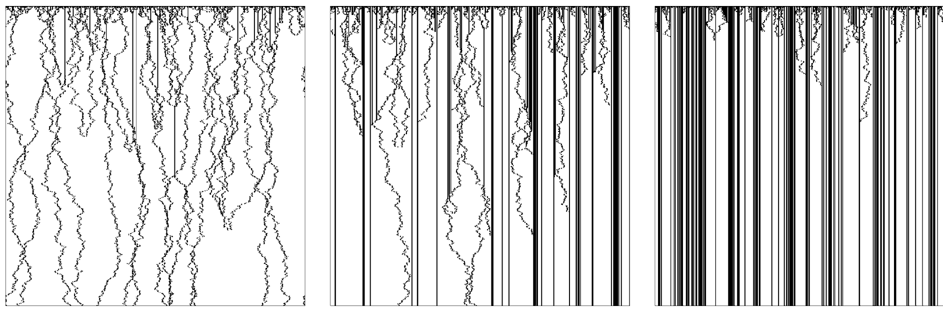
$$(u, i) \xrightarrow{a} (u + 1, i) \qquad \text{at time } t.$$

On the other hand, if the individuals at x and $x + 1$ still disagree on their ith feature after the update, then $(u + 1, i)$ is occupied at time t so both particles coalesce, which we write

$$(u, i) \xrightarrow{c} (u + 1, i) \qquad \text{at time } t.$$

We refer to Figure 2 for an illustration of the coupling between the four-feature, three-state Axelrod model and systems of annihilating-coalescing random walks. Each particle is represented by a cross and the three possible states by the colors black, grey and white. In our example, there are two jumps resulting in two collisions: an annihilating event then a coalescing event. We also refer the reader to Figure 3 for simulation pictures of the systems of random walks when $F = 3$.

Lanchier and Schweinsberg [9] observed that, when $q = 2$, random walks can only annihilate, which was the key to proving clustering. This is due to the fact that, in a simplistic world where there are only two possible alternatives for each cultural feature, two individuals who disagree with a third one must agree. In our context, the individuals at $x - 1$ and $x + 1$ must agree just before the update when $q = 2$, which results in an annihilating event. In contrast, when the number of states per feature is larger, the three consecutive vertices may have three different views on their ith cultural feature, which results in a coalescing event. We point out that, since the system of random walks collects all the times at which pairs of neighbors interact, the knowledge of the initial configuration of the Axelrod



(a) $F = 3$ and $q = 2$ (b) $F = 3$ and $q = 3$ (c) $F = 3$ and $q = 4$

FIG. 3. *System of annihilating-coalescing random walks on the torus with* 600 *vertices*.

model and the system of random walks up to time t allows us to re-construct the Axelrod model up to time t regardless of the value of the parameters. There is, however, a crucial difference depending on the number of states. When $q = 2$, collisions always result in annihilating events, so knowing the configuration of the Axelrod model is unimportant in determining the evolution of the random walks. In contrast, when $q > 2$, whether a collision results in a coalescing or an annihilating event depends on the configuration of the Axelrod model just before the time of the collision. The key to all our results is that, in spite of this dependency, collisions result independently in either an annihilating event or a coalescing event with some fixed probabilities. In particular, the outcome of a collision is independent of the past of the system of random walks though it is not independent of the past of the Axelrod model itself.

To prove this result, we need to construct the one-dimensional process graphically from a percolation structure and then define active paths which basically keep track of the descendants of the initial opinions. First, we consider the following collections of independent Poisson processes and random variables: for each pair of vertex and feature $(x, i) \in \mathbb{Z} \times \{1, 2, \dots, F\}$:

- we let $(N_{x,i}(t) : t \geq 0)$ be a rate one Poisson process;
- we denote by $T_{x,i}(n)$ its nth arrival time: $T_{x,i}(n) := \inf\{t : N_{x,i}(t) = n\}$;
- we let $(B_{x,i}(n) : n \geq 1)$ be a collection of independent Bernoulli variables with

$$P\big(B_{x,i}(n) = +1\big) = P\big(B_{x,i}(n) = -1\big) = 1/2;$$

- and we let $(U_{x,i}(n) : n \geq 1)$ be a collection of independent Uniform(0, 1).

The Axelrod model is then constructed as follows. At time $t = T_{x,i}(n)$, we draw an arrow labeled i from vertex x to vertex $y := x + B_{x,i}(n)$ to indicate that if

$$U_{x,i}(n) \leq r\big(\zeta_{t-}(u)\big) \quad \text{and} \quad \zeta_{t-}(u) \neq 0 \qquad \text{where } u = \frac{x+y}{2} \in \mathbb{D}, \tag{7}$$

then the individual at vertex y imitates the ith feature of the individual at vertex x. In particular, as indicated in (6), the rate at which the imitation occurs is equal to one half times the fraction of cultural features both vertices have in common divided by the number of features for which both vertices disagree, which indeed produces the local transition rates of the Axelrod model. The graphical representation defines a random graph structure, also called percolation structure, from which the process starting from any initial configuration can be constructed by induction based on an argument due to Harris [6]. Each arrow in this percolation structure is said to be active if condition (7) is satisfied. Note that whether an arrow is active or not depends on the initial configuration, and that the fact that an i-arrow from vertex x to vertex y at time t is active implies that the ith feature of y must be equal to the ith feature of x at time t. We say that there is an active i-path from (z, s) to (x, t) whenever there are sequences of times and vertices

$$s_0 = s < s_1 < \cdots < s_{n+1} = t \quad \text{and} \quad x_0 = z, x_1, \ldots, x_n = x$$

such that the following two conditions hold:

(1) For $j = 1, 2, \ldots, n$, there is an active i-arrow from x_{j-1} to x_j at time s_j.
(2) For $j = 0, 1, \ldots, n$, there is no active i-arrow that points at $\{x_j\} \times (s_j, s_{j+1})$.

We say that there is a generalized active path from (z, s) to (x, t) whenever

(3) for $j = 1, 2, \ldots, n$, there is an active arrow from x_{j-1} to x_j at time s_j.

Later, we will use the notation $\overset{i}{\rightsquigarrow}$ and $\rightsquigarrow$ to indicate the existence of an active i-path and a generalized active path, respectively. Conditions 1 and 2 above imply that

$$\text{for all } (x, t) \in \mathbb{Z} \times \mathbb{R}_+ \text{ and all } i, \text{ there is a unique } z \in \mathbb{Z} \text{ such that } (z, 0) \overset{i}{\rightsquigarrow} (x, t).$$

Moreover, because of the definition of active arrows and simple induction, the ith cultural feature of vertex x at time t is equal to the initial value of the ith cultural feature of z, so we call vertex z the ancestor of vertex x at time t for the ith feature. In contrast, generalized active paths, which can be seen as concatenations of active i-paths for possibly different values of i, do not have such an interpretation, but the concept will be useful later to prove fixation.

LEMMA 3. *Conditional on the realization of the system of random walks until time $t-$ and the event that $(u, i) \longrightarrow (u + 1, i)$ at time t, we have*

$$(u, i) \xrightarrow{a} (u + 1, i) \text{ at time } t \text{ with probability } (q - 1)^{-1}$$
$$(u, i) \xrightarrow{c} (u + 1, i) \text{ at time } t \text{ with probability } (q - 2) \cdot (q - 1)^{-1}.$$

PROOF. Let $x := u + 1/2 \in \mathbb{Z}$. Due to one-dimensional nearest neighbor interactions, active i-paths cannot cross each other, from which we deduce that

$$a_s(x - 1, i) \le a_s(x, i) \le a_s(x + 1, i) \qquad \text{for all } s \ge 0, \tag{8}$$

where $a_s(\cdot, i)$ denotes the ancestor at time s for the ith feature, that is,

$$\big(a_s(y, i), 0\big) \overset{i}{\rightsquigarrow} (y, s) \qquad \text{for } y \in \{x-1, x, x+1\} \text{ and all } s \geq 0.$$

Moreover, conditional on the event of a collision $(u, i) \longrightarrow (u+1, i)$ at time t, there is a particle at (u, i) and a particle at $(u+1, i)$ at time $t-$, therefore

$$\eta_0\big(a_{t-}(x \pm 1, i)\big) = \eta_{t-}(x \pm 1, i) \neq \eta_{t-}(x, i) = \eta_0\big(a_{t-}(x, i)\big). \tag{9}$$

From (8) and (9), we deduce that, conditional on $(u, i) \longrightarrow (u+1, i)$ at time t,

$$a_s(x-1, i) < a_s(x, i) < a_s(x+1, i) \qquad \text{for all } s < t.$$

In other respects, we have

$$\begin{aligned} &(u, i) \overset{a}{\longrightarrow} (u+1, i) \text{ at time } t \\ &\qquad \text{if and only if } (u, i) \longrightarrow (u+1, i) \text{ at time } t \text{ and} \\ &\qquad\quad \eta_{t-}(x-1, i) = \eta_{t-}(x+1, i), \\ &\qquad \text{if and only if } (u, i) \longrightarrow (u+1, i) \text{ at time } t \text{ and} \\ &\qquad\quad \eta_0\big(a_{t-}(x-1, i)\big) = \eta_0\big(a_{t-}(x+1, i)\big). \end{aligned}$$

In particular, the outcome—either an annihilating event or a coalescing event—of a collision at time t is independent of the realization of the system of random walks up to time $t-$. Moreover, since the initial states are independent and uniformly distributed, the conditional probability of an annihilating event is equal to the conditional probability

$$P(X = Z | X \neq Y \text{ and } Z \neq Y), \tag{10}$$

where X, Y, Z are independent uniform random variables over $\{1, 2, \ldots, q\}$. By conditioning on the possible values of Y, we obtain that (10) is equal to

$$\sum_{j=1}^{q} P(X = Z | X \neq j \text{ and } Z \neq j) P(Y = j) = \sum_{j=1}^{q} \big((q-1)q\big)^{-1} = (q-1)^{-1}.$$

Finally, since each collision results in either an annihilating event or a coalescing event, the conditional probability of a coalescing event directly follows. This completes the proof. □

3. Sufficient condition for fixation. The main objective of this section is to extend a result of [4] to the Axelrod model, and obtain a sufficient condition for fixation which is based on certain properties of the active i-paths.

LEMMA 4. *For all* $(z, i) \in \mathbb{Z} \times \{1, 2, \ldots, F\}$, *let*

$$T(z, i) := \inf\big\{t : (z, 0) \overset{i}{\rightsquigarrow} (0, t)\big\}.$$

Then, the Axelrod model fixates whenever

(11) $$\lim_{N\to\infty} P\big(T(z,i) < \infty \text{ for some } z < -N \text{ and some } i = 1, 2, \dots, F\big) = 0.$$

PROOF. Extending an idea of Bramson and Griffeath [4] and generalizing the technique in [10], we set $\tau_{i,0} := 0$ for every cultural feature i and define recursively the sequence of stopping times

$$\tau_{i,j} := \inf\{t > \tau_{i,j-1} : \eta_t(0,i) \neq \eta_{\tau_{i,j-1}}(0,i)\} \qquad \text{for } j \geq 1.$$

In other words, the stopping time $\tau_{i,j}$ is the jth time the individual at the origin changes the state of her ith cultural feature. Also, for each cultural feature i, we define the random variables

$$a_{i,j} := \text{the ancestor of vertex 0 at time } \tau_{i,j} \text{ for the } i\text{th feature}$$

as well as the collection of events

$$B_i := \{\tau_{i,j} < \infty \text{ for all } j\} \quad \text{and} \quad G_{i,N} := \{|a_{i,j}| < N \text{ for all } j\}.$$

See the left-hand side of Figure 4 for a schematic illustration of the stopping times $\tau_{i,j}$ and the corresponding vertices $a_{i,j}$. Assumption (11) together with reflection symmetry implies that, for each cultural feature i, the event $G_{i,N}$ occurs almost surely for some N. It follows that

$$P\left(\bigcup_{i=1}^F B_i\right) \leq \sum_{i=1}^F P(B_i) = \sum_{i=1}^F P\left(B_i \cap \left(\bigcup_N G_{i,N}\right)\right) = \sum_{i=1}^F P\left(\bigcup_N (B_i \cap G_{i,N})\right).$$

Since the event that the individual at the origin changes her culture infinitely often is also the event that at least one of the events B_i occurs, in view of the previous inequality, in order to establish fixation, it suffices to prove that

(12) $$P(B_i \cap G_{i,N}) = 0 \qquad \text{for all } i \in \{1, 2, \dots, F\} \text{ and all } N \geq 1.$$

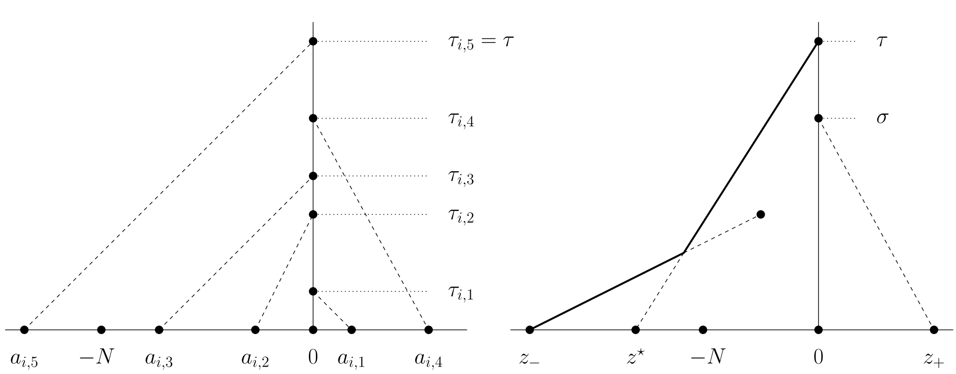


FIG. 4. *Picture related to the proof of Lemma 4. Dashed lines represent active i-paths for some i whereas the continuous thick line on the right-hand side is a generalized active path as defined in Section 2.*

Our proof of (12) relies on some symmetry properties of the Axelrod model that do not hold for the cyclic particle systems considered in [4]. First, we let

$$I_t(x,i) := \{z \in \mathbb{Z} : (x,i) \text{ is the ancestor of } (z,i) \text{ at time } t\}$$

be the set of descendants of (x,i) at time t, and denote by $M_t(x,i)$ its cardinality. Since each interaction between two individuals is equally likely to affect the culture of each of these two individuals, the number of descendants of any given site is a martingale whose expected value is constantly equal to one. In particular, the martingale convergence theorem implies that

$$\lim_{t\to\infty} M_t(x,i) = M_\infty(x,i) \qquad \text{with probability 1 where } E|M_\infty(x,i)| < \infty.$$

Therefore, for almost all realizations of the process, the number of descendants of (x,i) converges to a finite value. Since in addition the number of descendants is an integer-valued process,

$$\sigma(x,i) := \inf\{t > 0 : M_t(x,i) = M_\infty(x,i)\} < \infty \qquad \text{with probability 1.}$$

Using that simultaneous updates occur with probability zero, we deduce that the set of descendants inherits the properties of its cardinality in the sense that, with probability one,

$$\begin{aligned} &\lim_{t\to\infty} I_t(x,i) = I_\infty(x,i) \quad \text{and} \\ &\rho(x,i) := \inf\{t > 0 : I_t(x,i) = I_\infty(x,i)\} < \infty, \end{aligned} \tag{13}$$

where, due to one-dimensional nearest neighbor interactions, $I_\infty(x,i)$ is a random interval which is almost surely finite. To conclude, we simply observe that, conditional on $G_{i,N}$, the last time the individual at the origin changes the state of her ith cultural feature is at most equal to the largest of the stopping times $\rho(x,i)$ for $x \in (-N,N)$ from which it follows that

$$P(B_i \cap G_{i,N}) = P(\rho(x,i) = \infty \text{ for some } -N < x < N) = 0$$

according to (13). This proves (12) and therefore the lemma. $\square$

4. Proof of Theorem 1. In view of Lemma 4, in order to prove fixation, it suffices to show that the probability of the event in equation (11), that we denote by H_N, tends to zero as $N \to \infty$. The first step is to extend the construction proposed by Bramson and Griffeath [4] to the Axelrod model, the main difficulty being that two active paths at different levels can cross each other. Let τ be the first time an active i-path for some $i = 1, 2, \ldots, F$ that originates from $(-\infty, -N)$ hits the origin, and observe that

$$\tau = \inf\{T(z,i) : z \in (-\infty,-N) \text{ and } i = 1, 2, \ldots, F\}$$

from which it follows that

$$H_N := \{T(z,i) < \infty \text{ for some } (z,i) \in (-\infty,-N) \times \{1,2,\ldots,F\}\} = \{\tau < \infty\}.$$

Denote by $z^\star < -N$ the initial position of this active path. Also, we set

$$(14)\qquad \begin{aligned} z_- &:= \min\{z \in \mathbb{Z} : (z,0) \rightsquigarrow (0,\tau)\} \leq z^\star < -N, \\ z_+ &:= \max\{z \in \mathbb{Z} : (z,0) \rightsquigarrow (0,\sigma) \text{ for some } \sigma < \tau\} \geq 0 \end{aligned}$$

and define $I = (z_-, z_+)$. We point out that $z_- < z^\star$ in general since vertex $z^\star$ is defined from the set of active i-paths whereas vertex z_- is defined from generalized active paths that are concatenations of active i-paths with different values of i. See the right-hand side of Figure 4 for an illustration where the two vertices are different. Now, note that each blockade which is initially in the interval I must have been destroyed, that is, turned into a set of $F-1$ active particles through the annihilation of one of the particles that constitute the blockade, by time τ. Moreover, active particles initially outside the interval I cannot jump inside the space–time region delimited by the two generalized active paths implicitly defined in (14). Indeed, assuming that such particles exist would contradict either the minimality of z_- or the maximality of z_+. In particular, on the event H_N, all the blockades initially in I must have been destroyed before time τ by either active particles initially in I or active particles that result from these blockade destructions. To estimate the probability of this last event, we first give a weight of -1 to each particle initially active by setting

$$\phi(u) := -\zeta_0(u) = -i \qquad \text{whenever } \zeta_0(u) = i \neq F.$$

To define $\phi(u)$ when u is initially occupied by a blockade, we observe that by Lemma 3 the number of collisions required to break a blockade is geometric with mean $q-1$. Moreover, each blockade destruction results in a total of $F-1$ active particles. Therefore, we set

$$\phi(u) := \psi(u) - (F-1) \qquad \text{whenever } \zeta_0(u) = F,$$

where $\psi(u)$ are independent geometric random variables with mean $q-1$. The fact that H_N occurs only if all the blockades initially in I are destroyed by active particles initially in I or active particles resulting from these blockade destructions, can then be written as

$$(15)\qquad \begin{aligned} H_N &\subset \left\{\sum_{u \in I} \phi(u) \leq 0\right\} \\ &\subset \left\{\sum_{u=l}^{r} \phi(u) \leq 0 \text{ for some } l < -N \text{ and some } r \geq 0\right\}. \end{aligned}$$

To understand the first inclusion, simply observe that the sum of the $\phi(u)$ is equal to the number of collisions required to break all the blockades minus the total number of active particles initially in the interval I or created from the destruction of blockades initially in I. Since the number of collisions is bounded by the number of such active particles, all the blockades initially in I can only be destroyed if the number of such active particles exceeds the number of collisions required, which gives the first inclusion. The second inclusion simply follows from the fact that

$$(-N,0) \subset (z_-, z_+) = I \qquad \text{since } z_- < -N \text{ and } z_+ \geq 0.$$

The expression of $\omega(q,F)$ can be understood heuristically as follows: since the $\phi(u)$ are independent, one expects that fixation occurs if $E\phi(u) > 0$. But

$$\begin{aligned} E\phi(u) &= \big(E\psi(u) - (F-1)\big) P\big(\zeta_0(u) = F\big) - \sum_{i=0}^{F-1} i\, P\big(\zeta_0(u) = i\big) \\ &= \big(E\psi(u) + 1\big) P\big(\zeta_0(u) = F\big) - \sum_{i=0}^{F} i\, P\big(\zeta_0(u) = i\big) \\ &= q\, P\big(\zeta_0(u) = F\big) - E\zeta_0(u), \end{aligned}$$

which, since $\zeta_0(u) = \text{Binomial}(F, 1 - 1/q)$, is precisely equal to $\omega(q,F)$. To deduce rigorously fixation from the positiveness of the expected value, which is done in the next two lemmas, we now prove large deviation estimates for H_N. The first of these two lemmas will be used in the proof of the second one to show that the total number of collisions required to break all the blockades in a large interval does not deviate too much from its expected value.

LEMMA 5. *Let $X_1, X_2, \ldots$ be an infinite sequence of independent geometric random variables with the same parameter p. Then, for all $\varepsilon > 0$, there exists $\gamma_1 > 0$ such that*

$$P\big(X_1 + X_2 + \cdots + X_K \leq (1/p - \varepsilon)K\big) \leq \exp(-\gamma_1 K) \qquad \textit{for all } K \textit{ sufficiently large}.$$

PROOF. Let $Z_n = \text{Binomial}(n, p)$ for all $n \geq 1$. Since, in a sequence of independent Bernoulli trials with success probability p, the event that the Kth success occurs at step n is included in the event that K successes occur in the first n steps, we have

$$P(X_1 + X_2 + \cdots + X_K = n) \leq P(Z_n = K).$$

Letting M denote the integer part of $(1/p - \varepsilon)K$, we deduce that

$$\begin{aligned} &P\big(X_1 + X_2 + \cdots + X_K \leq (1/p - \varepsilon)K\big) \\ &\quad = \sum_{n=K}^{M} P(X_1 + X_2 + \cdots + X_K = n) \\ &\quad \leq \sum_{n=K}^{M} P(Z_n = K) \leq \sum_{n=K}^{M} P(Z_n \geq K) \leq \sum_{n=K}^{M} P(Z_M \geq K) \\ &\quad \leq M \times P(Z_M \geq K). \end{aligned}$$

Since large deviation estimates for the binomial distribution imply that

$$P(Z_M \geq K) \leq P\big(Z_M \geq (1-\varepsilon p)^{-1} Mp\big) \\ \leq \exp(-\gamma_2 M) \leq \exp\big(-\gamma_2\big((1/p-\varepsilon)K - 1\big)\big)$$

for a suitable constant $\gamma_2 > 0$, the result follows. □

LEMMA 6. *Let $I_N := (-N, 0) \cap \mathbb{D}$ and assume that $\omega(q, F) > 0$. Then*

$$P\bigg(\sum_{u \in I_N} \phi(u) \leq 0\bigg) \leq \exp(-\gamma_3 N)$$

for a suitable constant $\gamma_3 > 0$ and all N sufficiently large.

PROOF. To begin with, we define

$$N_i := \operatorname{card}\{u \in I_N : \zeta_0(u) = i\} \qquad \text{for } i = 0, 1, \dots, F.$$

Since the random variables $\zeta_0(u)$, $u \in \mathbb{D}$, are independent, standard large deviation estimates for the binomial distribution imply that for all $\varepsilon > 0$ there exists $\gamma_4 > 0$ such that

$$P\big(N_i \notin \big((\mu_i - \varepsilon)N, (\mu_i + \varepsilon)N\big)\big) \leq \exp(-\gamma_4 N) \qquad \text{for all } i = 0, 1, \dots, F, \tag{16}$$

where $\mu_i := P(X = i)$ with $X = \text{Binomial}(F, 1 - 1/q)$. The expression for μ_i follows from the fact that initially each level of each site is independently occupied with probability $1 - 1/q$, which implies that the $\zeta_0(u)$ are independent binomial random variables. Let Ω be the event that

$$(\mu_i - \varepsilon)N < N_i < (\mu_i + \varepsilon)Na \qquad \text{for all } i = 0, 1, \dots, F.$$

Then, there exists a constant $C > 0$ such that, on the event Ω,

$$\frac{1}{N}\sum_{i=0}^{F-1} i N_i \leq \sum_{i=0}^{F-1} i(\mu_i + \varepsilon) \leq \sum_{i=0}^{F} i\mu_i - F\mu_F + C\varepsilon = E\zeta_0(u) - F\mu_F + C\varepsilon.$$

In particular, letting K be the integer part of $(\mu_F - \varepsilon)N$, we have

$$\begin{aligned} & P\bigg(\sum_{u \in I_N} \phi(u) \leq 0 \,\bigg|\, \Omega\bigg) \\ & \quad \leq P\bigg(\sum_{u \in I_K} \big(\psi(u) - (F-1)\big) \leq \big(E\zeta_0(u) - F\mu_F + C\varepsilon\big)N\bigg) \\ & \quad \leq P\bigg(\sum_{u \in I_K} \psi(u) \leq \big(E\zeta_0(u) - \mu_F + (C - F + 1)\varepsilon\big)N\bigg). \end{aligned} \tag{17}$$

Now, since $\omega(q, F) > 0$, there exists $\varepsilon > 0$ small such that

$$\begin{aligned} E\zeta_0(u) - \mu_F + (C - F + 1)\varepsilon &= (q-1)\mu_F + E\zeta_0(u) - q\mu_F + (C - F + 1)\varepsilon \\ &= (q-1)\mu_F - \omega(q, F) + (C - F + 1)\varepsilon \\ &\leq (q - 1 - \varepsilon)(\mu_F - \varepsilon) \end{aligned}$$

from which we deduce, also using (17) and Lemma 5, that

$$P\biggl(\sum_{u \in I_N} \phi(u) \leq 0 \Big| \Omega\biggr) \leq P\biggl(\sum_{u \in I_K} \psi(u) \leq (q - 1 - \varepsilon)K\biggr) \leq \exp(-\gamma_1 K) \tag{18}$$

for all K sufficiently large. Combining (16) and (18), we obtain

$$P\biggl(\sum_{u \in I_N} \phi(u) \leq 0\biggr) \leq \exp\bigl(-\gamma_1(\mu_F - \varepsilon)N\bigr) + (F + 1)\exp(-\gamma_4 N)$$

for all N sufficiently large. $\square$

Using the inclusion in (15) and Lemma 6, we deduce

$$\begin{aligned} \lim_{N\to\infty} P(H_N) &\leq \lim_{N\to\infty} P\biggl(\sum_{u=l}^{r} \phi(u) \leq 0 \text{ for some } l < -N \text{ and some } r \geq 0\biggr) \\ &\leq \lim_{N\to\infty} \sum_{l<-N} \sum_{r\geq 0} P\biggl(\sum_{u=l}^{r} \phi(u) \leq 0\biggr) \\ &\leq \lim_{N\to\infty} \sum_{l<-N} \sum_{r\geq 0} \exp\bigl(-\gamma_3(r - l)\bigr) = 0. \end{aligned}$$

This, together with Lemma 4, implies fixation whenever $\omega(q, F) > 0$.

5. Fixation when $F = 2$ and $q = 3$. To begin with, note that, when $F = 2$ and $q = 3$, we have $E\phi(u) = \omega(3, 2) = 0$ for the comparison function $\phi(u)$ defined in the previous section. In particular, to find a good enough upper bound for the probability of H_N in the case $F = 2$ and $q = 3$, one needs to define a new comparison function that also takes into account additional events that promote fixation, such as collisions between active particles and blockade formations. Recall that in the comparison function of Section 4, each particle which is initially active is assigned a weight of -1, which corresponds to the worst case scenario in which the active particle hits a blockade. However, each active particle can also hit another active particle or form a new blockade with another active particle. More precisely, there are four possible outcomes for each active particle:

(1) If the active particle hits a blockade, it is assigned a weight of -1.

(2) If the active particle coalesces with another active particle, then at most one collision with a blockade can result from this pair of particles so the pair is assigned a total weight of -1; that is, each particle of the pair is individually assigned a weight of $-1/2$.

(3) If the active particle annihilates with another active particle, then no collision with a blockade can result from this pair so each active particle that annihilates with another active particle is assigned a weight of 0.

(4) If the active particle forms a blockade with another active particle, then following the same approach as in the previous section the pair is assigned a total weight equal to -1 plus a geometric random variable with mean $q-1$.

In view of cases 2–4 above, the weight of an active particle that either collides with another active particle or forms a blockade with another active particle is at least $-1/2$, and therefore we define a new comparison function, again denoted by ϕ, as follows:

$$\phi(u) := \begin{cases} \psi(u) - 1, & \text{if } \zeta_0(u) = 2, \\ 0, & \text{if } \zeta_0(u) = 0, \\ -1/2, & \text{if } \zeta_0(u) = 1 \text{ and the active particle initially at } u \text{ either} \\ & \text{collides with another active particle or forms} \\ & \text{a blockade with another active particle,} \\ -1, & \text{if } \zeta_0(u) = 1 \text{ and the active particle initially at } u \\ & \text{collides with a blockade,} \end{cases}$$

where the random variables $\psi(u)$ are again independent geometric random variables with the same expected value $q-1=2$. The value of $\phi(u)$ when $\zeta_0(u) \neq 1$ is the same as in the previous section whereas we distinguish between active particles that satisfy case 1 or cases 2–4 above. The same reasoning and construction as in Section 4 again imply that

$$H_N \subset \left\{ \sum_{u=l}^{r} \phi(u) \leq 0 \text{ for some } l < -N \text{ and some } r \geq 0 \right\} \tag{19}$$

for this new comparison function. To prove that the probability of the event on the right-hand side converges to zero as $N \to \infty$, we follow the same strategy as for Lemma 6 but also find a lower bound for the probability that a particle initially active either collides with another active particle or forms a blockade with another active particle, which is done in the next lemma.

LEMMA 7. *Assume that $F = 2$ and $q \geq 3$. Then, there exists $\gamma_5 > 0$ such that*

$$P\left(\sum_{u \in I_N} \phi(u) \leq 0 \right) \leq \exp(-\gamma_5 N) \qquad \textit{for all } N \textit{ sufficiently large},$$

where $I_N := (-N, 0) \cap \mathbb{D}$ as in Lemma 6.

PROOF. The first step is to find a lower bound for the initial number of active particles that will either collide or form a blockade with another active particle. To do so, we introduce the following definition: an active particle initially at site $u \in \mathbb{D}$ is said to be a good particle if

$$\zeta_0(u) = \zeta_0(v) = 1 \quad \text{where } \{u, v\} = \{2n - 1/2, 2n + 1/2\} \text{ for some } n \in \mathbb{Z}. \tag{20}$$

In other words, we partition the lattice $\mathbb{D}$ into countably many pairs of adjacent sites, and call an active particle at time 0 a good particle if the other site of the pair is initially occupied by an active particle as well. An active particle which is not good is called a bad particle. Since initially each level of each site is independently occupied with probability $1 - 1/q$, the variables $\zeta_0(u)$ are independent binomial random variables, so for u, v as in (20) we have

$$P(\{u, v\} \text{ is occupied by a pair of good particles at time } 0) = \nu_0 = P(X = 1)^2,$$

where $X = \text{Binomial}(2, 1 - 1/q)$. Similarly, we have

$$P(u \text{ is occupied by a bad particle at time } 0) = \nu_1 = P(X = 1) \times P(X \neq 1),$$

$$P(u \text{ is occupied by two particles at time } 0) = \nu_2 = P(X = 2).$$

Since in addition the events that nonoverlapping pairs of adjacent sites are initially occupied by two good particles, or one bad particle, or one blockade, or one bad particle and one blockade, or two blockades are independent, standard large deviation estimates for the binomial distribution imply that there exists a positive constant $\gamma_6 > 0$ such that

$$P\big(N_i \notin ((\nu_i - \varepsilon)N, (\nu_i + \varepsilon)N)\big) \leq \exp(-\gamma_6 N) \qquad \text{for } i = 0, 1, 2, \tag{21}$$

where N_0, N_1 and N_2 denote respectively the initial number of good particles, the initial number of bad particles and the initial number of blockades in the interval I_N. To estimate the probability that a pair of good particles collide or form a blockade, we first observe that, when there are only two features, the graphical representation of the Axelrod model simplifies as follows: For each pair of neighbors $(x, y) \in \mathbb{Z}^2$, draw an arrow $x \to y$ at the times of a Poisson process with intensity one fourth, which is equal to half of the rate at which neighbors who agree on one cultural feature interact. If the two neighbors agree on exactly one cultural feature at the time of the interaction then the culture of the individual at vertex y becomes the same as the culture of the individual at vertex x. In this graphical representation, there are exactly six possible arrows that may affect the system of random walks at the pair of sites $\{u, u + 1\} \subset \mathbb{D}$, namely

$$\begin{aligned} &u - 1/2 \to u + 1/2, \qquad && u + 3/2 \to u + 1/2, \\ &u + 1/2 \to u - 1/2, \qquad && u + 1/2 \to u + 3/2, \\ &u - 3/2 \to u - 1/2, \qquad && u + 5/2 \to u + 3/2. \end{aligned} \tag{22}$$

The event that one of the two arrows in the first line of (22) appears before any of the four other ones occurs with probability two (arrows) over six (arrows) = 1/3, and on the intersection of this event and the event that there is initially a pair of good particles at $\{u, u+1\}$, the two particles either collide or form a blockade. Moreover, the event that one of the two arrows in the first line appears first only depends on the realization of the graphical representation in

$$(u - 3/2, u + 5/2) \times [0, \infty).$$

In particular, parts of the graphical representation associated with nonadjacent pairs do not intersect which, by independence of the Poisson processes, implies that the events that the two arrows in the first line of (22) appears before any of the other ones are independent for nonadjacent pairs. It follows that the initial number J of good particles in I_N that either collide or form a blockade is stochastically larger than a binomial random variable with $N\nu_0/2$ trials and success probability one third. Large deviation estimates for the binomial distribution then imply that

$$P\big(J \le (1/6 - \varepsilon)(\nu_0 - \varepsilon)N \,|\, N_0 > (\nu_0 - \varepsilon)N\big) \le \exp(-\gamma_7 N) \tag{23}$$

for a suitable constant $\gamma_7 > 0$. Now, let Ω be the event that

$$(\nu_i - \varepsilon)N < N_i < (\nu_i + \varepsilon)N \qquad \text{for } i = 0, 1, 2 \quad \text{and} \quad J > (1/6 - \varepsilon)(\nu_0 - \varepsilon)N,$$

and observe that there exists a constant $C > 0$ such that, on the event Ω,

$$\begin{aligned}(1/2)J + (N_0 + N_1 - J) &= N_0 + N_1 - (1/2)J \\ &< (\nu_0 + \nu_1 + 2\varepsilon)N - (1/2)(1/6 - \varepsilon)(\nu_0 - \varepsilon)N \\ &= (11\nu_0/12 + \nu_1 + C\varepsilon)N.\end{aligned}$$

In particular, letting K be the integer part of $(\nu_2 - \varepsilon)N$, we have

$$\begin{aligned}&P\Big(\sum_{u \in I_N} \phi(u) \le 0 \Big| \Omega\Big) \\ &\quad \le P\Big(\sum_{u \in I_K} \big(\psi(u) - 1\big) \le (11\nu_0/12 + \nu_1 + C\varepsilon)N\Big) \\ &\quad \le P\Big(\sum_{u \in I_K} \psi(u) \le \big(11\nu_0/12 + \nu_1 + \nu_2 + (C-1)\varepsilon\big)N\Big).\end{aligned} \tag{24}$$

In other respects, recalling the definition of ν_i for $i = 0, 1, 2$, we have

$$\begin{aligned}&(q-2)\nu_2 - \nu_1 - 11\nu_0/12 \\ &\quad = (q-2)P(X=2) - P(X=1)P(X \ne 1) - (11/12)P(X=1)^2 \\ &\quad = (q-2)P(X=2) - P(X=1) + (1/12)P(X=1)^2,\end{aligned}$$

which, recalling the definition of X, is equal to

$$(q-2)\left(1-\frac{1}{q}\right)^2-\frac{2}{q}\left(1-\frac{1}{q}\right)+\frac{1}{12}\left(\frac{2}{q}\left(1-\frac{1}{q}\right)\right)^2$$
$$=(q-3)\left(1-\frac{1}{q}\right)+\frac{1}{3}\left(\frac{1}{q}\left(1-\frac{1}{q}\right)\right)^2\geq\frac{1}{3}\left(\frac{1}{3}\left(1-\frac{1}{3}\right)\right)^2=\frac{4}{243}>0$$

for all $q\geq 3$. In particular, there exists $\varepsilon>0$ small such that

$$\begin{aligned}
&11\nu_0/12+\nu_1+\nu_2+(C-1)\varepsilon\\
&\quad=(q-1)\nu_2-\big((q-2)\nu_2-\nu_1-11\nu_0/12\big)+(C-1)\varepsilon\\
&\quad=(q-1)\nu_2-(\nu_2+q-1)\varepsilon\leq(q-1-\varepsilon)(\nu_2-\varepsilon).
\end{aligned}$$

Since $E\psi(u)=q-1$, the previous estimate, (24) and Lemma 5 imply that

$$P\left(\sum_{u\in I_N}\phi(u)\leq 0\middle|\Omega\right)\leq P\left(\sum_{u\in I_K}\psi(u)\leq(q-1-\varepsilon)K\right)\leq\exp(-\gamma_1 K) \tag{25}$$

for all K sufficiently large. Combining (21), (23) and (25), we obtain

$$P\left(\sum_{u\in I_N}\phi(u)\leq 0\right)\leq\exp\big(-\gamma_1(\nu_2-\varepsilon)N\big)+3\exp(-\gamma_6 N)+\exp(-\gamma_7 N)$$

for all N sufficiently large, which completes the proof. □

As in the previous section, (19) and Lemma 7 imply that

$$\lim_{N\to\infty}P(H_N)\leq\lim_{N\to\infty}\sum_{l<-N}\sum_{r\geq 0}\exp\big(-\gamma_5(r-l)\big)=0,$$

which, together with Lemma 4, implies fixation when $F=2$ and $q=3$.

Acknowledgments. The authors would like to thank an anonymous referee for her/his careful reading of the proofs and suggestions to improve the clarity of the paper, and for pointing out a mistake in a preliminary version of the proof of Lemma 7.

REFERENCES

[1] AXELROD, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *J. Conflict. Resolut.* **41** 203–226.

[2] BRAMSON, M. and GRIFFEATH, D. (1980). On the Williams–Bjerknes tumour growth model. II. *Math. Proc. Cambridge Philos. Soc.* **88** 339–357. MR0578279

[3] BRAMSON, M. and GRIFFEATH, D. (1981). On the Williams–Bjerknes tumour growth model. I. *Ann. Probab.* **9** 173–185. MR0606980

[4] BRAMSON, M. and GRIFFEATH, D. (1989). Flux and fixation in cyclic particle systems. *Ann. Probab.* **17** 26–45. MR0972768

[5] CLIFFORD, P. and SUDBURY, A. (1973). A model for spatial conflict. *Biometrika* **60** 581–588. MR0343950

[6] HARRIS, T. E. (1972). Nearest-neighbor Markov interaction processes on multidimensional lattices. *Adv. Math.* **9** 66–89. MR0307392

[7] HOLLEY, R. A. and LIGGETT, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3** 643–663. MR0402985

[8] LANCHIER, N. (2012). The Axelrod model for the dissemination of culture revisited. *Ann. Appl. Probab.* **22** 860–880. MR2953571

[9] LANCHIER, N. and SCHWEINSBERG, J. (2012). Consensus in the two-state Axelrod model. *Stochastic Process. Appl.* **122** 3701–3717. MR2965921

[10] SCARLATOS, S. (2012). Behavior of social dynamical models I: Fixation in the symmetric cyclic system (with paradoxical effect in the six-color automaton). In *ACRI* 2012. *LNCS*, (G. C. Sirakoulis and S. Bandini, eds.) **7495** 141–150. Springer, Heidelberg.

[11] VILONE, D., VESPIGNANI, A. and CASTELLANO, C. (2002). Ordering phase transition in the one-dimensional Axelrod model. *Eur. Phys. J. B* **30** 399–406.

SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES
ARIZONA STATE UNIVERSITY
TEMPE, ARIZONA 85287
USA

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF PATRAS
PATRAS 26500
GREECE