THE ROLE OF DISPERSAL IN INTERACTING PATCHES SUBJECT TO AN ALLEE EFFECT

N. LANCHIER,* *Arizona State University*

Abstract

This article is concerned with a stochastic multipatch model in which each local population is subject to a strong Allee effect. The model is obtained by using the framework of interacting particle systems to extend a stochastic two-patch model that was recently introduced by Kang and the author. The main objective is to understand the effect of the geometry of the network of interactions, which represents potential migrations between patches, on the long-term behavior of the metapopulation. In the limit as the number of patches tends to ∞, there is a critical value for the Allee threshold below which the metapopulation expands and above which the metapopulation goes extinct. Spatial simulations on large regular graphs suggest that this critical value strongly depends on the initial distribution when the degree of the network is large, whereas the critical value does not depend on the initial distribution when the degree is small. Looking at the system starting with a single occupied patch on the complete graph and on the ring, we prove analytical results that support this conjecture. From an ecological perspective, these results indicate that, upon arrival of an alien species subject to a strong Allee effect to a new area, though dispersal is necessary for its expansion, fast long-range dispersal drives the population toward extinction.

Keywords: Interacting particle system; Allee effect; dispersal; metapopulation

2010 Mathematics Subject Classification: Primary 60K35

1. Introduction

In population dynamics, the term Allee effect refers to a certain process that leads to decreasing net population growth with decreasing density. If, in addition, the growth rate becomes negative at low density, this monotone relationship results in the existence of a so-called Allee threshold below which populations are at high risk of being driven toward extinction, a phenomenon which is referred to as the strong Allee effect and may be due to various ecological factors: failure to locate mates, inbreeding depression, failure to satiate predators, lack of cooperative feeding, etc. The research on the causes and consequences of the presence of an Allee effect ranges from biological experiments to mathematical analyses of deterministic and stochastic models and is too copious to be reviewed in this paper. We refer the reader to [2] for a general overview of the literature on the Allee effect and also mention more particularly the recent work of Borrello [1], who, similarly to this paper, focused on a stochastic process based on the framework of interacting particle systems. His model is derived from a model proposed by Schinazi [7], [8] in which each site of the infinite regular lattice represents a patch that can

host a local population not exceeding a certain number of individuals, which can be thought of as a carrying capacity. In patches hosting a local population below the carrying capacity, individuals give birth at a fixed rate to offspring which stay in the parent's patch, whereas in patches at the carrying capacity, individuals give birth at another fixed rate to offspring which are sent to randomly chosen adjacent patches. The fact that offspring are placed preferentially in their parent's patch rather than adjacent patches models spatial aggregation, and the common value of the carrying capacity of the patches across the lattice can be seen as a measure of how much individuals aggregate. The analysis in [7] and [8] shows that, for appropriate values of the birth rates and in the presence of catastrophic events modeled by the death of each local population independently at rate 1, the metapopulation survives if and only if the maximum number of individuals per patch is smaller than some critical value. In contrast, in the absence of catastrophic events, which is modeled by the death of each individual independently at rate 1, the metapopulation survives if and only if the maximum number of individuals per patch is large enough. From an ecological perspective, these results suggest that, for species suffering local catastrophic events, spatial aggregation leads to extinction, whereas, for species suffering a strong Allee effect, spatial aggregation is necessary for survival. The model proposed by Borrello [1] is closely related to the previous model with individual deaths, but also includes an Allee threshold: there is a varying individual death rate that takes a larger value for local populations below the threshold, and a smaller value for local populations above the threshold and below the carrying capacity. The model allows, in addition, for the migration of flocks of individuals from local populations near the carrying capacity to adjacent patches, provided enough space is available. It is proved that, when only small flocks of individuals can migrate, the population goes extinct in the presence of a strong enough Allee effect, whereas, regardless of the strength of the Allee effect, survival is possible if large enough flocks can migrate, a phenomenon called the rescue effect in the context of metapopulations. The intuition behind this result is that small flocks of individuals moving to an empty patch are doomed to extinction, whereas the division of a local population into two subpopulations above the Allee threshold results in the successful invasion of an empty patch.

The main objective of this paper is to look more closely at the role of dispersal in populations subject to a strong Allee effect. To understand this aspect, Kang and the author recently introduced deterministic and stochastic two-patch models [6]. Here, we continue our analysis and use the framework of interacting particle systems to extend the stochastic two-patch model to a more general multipatch model. Thinking of the set of patches as the vertex set of a graph in which each edge indicates potential dyadic interactions, we study how the geometry of the network affects the survival probability of the metapopulation. In particular, instead of comparing models with different microscopic rules evolving on the same spatial structure as in the papers cited above, we compare models with the same microscopic rules but evolving on different spatial structures.

To define our multipatch model, we let $G = (V, E)$ be a graph representing the network of interactions. The system is a continuous-time Markov chain whose state at time t is a function

$$\eta_t \colon V \to [0, 1] \quad \text{with} \quad \eta_t(x) := \text{population density at vertex } x.$$

Having an Allee threshold $\theta \in (0, 1)$ and a migration factor $\mu \in (0, \frac{1}{2}]$, the evolution consists of the following two elementary events occurring in continuous time.

- *Mixing events.* Each edge becomes active at rate 1, which results in a fraction μ of the population at each of the two interacting vertices to move to the other vertex.

- *Local events.* Each vertex becomes active at rate 1, which results in the population density at that vertex to jump to state 0 if it is below the Allee threshold and to jump to 1 if it is above the Allee threshold. In case the density is exactly equal to the Allee threshold at the time of a local event, there is a jump to either state 0 or state 1 with probability $\frac{1}{2}$.

The inclusion of mixing events indicates that individuals can move from patch to patch through the edges of the graph. The strength of dispersal is thus modeled by the mean degree distribution of the network of interactions. Local events model the presence of a strong Allee effect in each patch: local populations below the Allee threshold are driven toward extinction whereas local populations above the Allee threshold expand, and we think of state 1 as the normalized density of a local population at the carrying capacity. In particular, our model includes the same two components as the model introduced in [1]: an Allee threshold and migrations between patches. Note however that migrations in our model are unconditional in the sense that there is no parameter that controls the size of the flocks moving between patches. This implies that, in contrast with [1], a mixing event may either increase or decrease the number of local populations above the Allee threshold and, therefore, be either beneficial or detrimental for the metapopulation. In particular, whether dispersal promotes survival or leads to extinction does not depend on the choice of the parameters but on the choice of the spatial structure. Our verbal description of the model in terms of local and mixing events can be expressed more formally through the Markov generator

$$
\begin{aligned}
L_\eta f(\eta) = &\sum_{(x,y) \in E} [f(\sigma_{x,y}\eta) - f(\eta)] \\
&+ \sum_{x \in V} \left(\mathbf{1}\{\eta(x) > \theta\} + \frac{1}{2} \mathbf{1}\{\eta(x) = \theta\} \right) [f(\sigma_x^+ \eta) - f(\eta)] \\
&+ \sum_{x \in V} \left(\mathbf{1}\{\eta(x) < \theta\} + \frac{1}{2} \mathbf{1}\{\eta(x) = \theta\} \right) [f(\sigma_x^- \eta) - f(\eta)], \quad\quad (1)
\end{aligned}
$$

where $(\sigma_x^+ \eta)(x) = 1$ and $(\sigma_x^- \eta)(x) = 0$, and

$$
(\sigma_{x,y}\eta)(z_1) = \eta(z_1) + \mu(\eta(z_2) - \eta(z_1)) \quad \text{whenever} \quad \{z_1, z_2\} = \{x, y\},
$$

while the state at all other vertices is unchanged. The stochastic two-patch model we introduced and studied in [6] is simply the process (1) when the network of interactions consists of two vertices connected by a single edge. The main objective there was to answer the following question: starting with one empty patch and one patch at the carrying capacity, does the inclusion of mixing events lead to a global extinction, i.e. both patches in state 0 eventually, or to a global expansion, i.e. both patches in state 1 eventually? Theorem 8 of [6] gave the following answer: when the migration factor is small, the probabilities of global extinction and global expansion are close to 1 when the Allee threshold is larger than $\frac{1}{2}$ and, respectively, smaller than $\frac{1}{2}$. As the migration factor increases, the limit becomes less predictable. This result suggests that $\frac{1}{2}$ is a critical value for the Allee threshold. However, we literally interpreted this $\frac{1}{2}$ as one patch initially in state 1 divided by two patches, and also conjectured that, for systems in which all the patches are connected, i.e. the network of interactions is a complete graph, the critical value for the Allee threshold must be equal to the initial fraction of patches in state 1 in the limit as the size of the system tends to ∞. This conjecture is supported by the bottom-right-hand picture of Figure 1 which shows numerical results for the stochastic process (1)
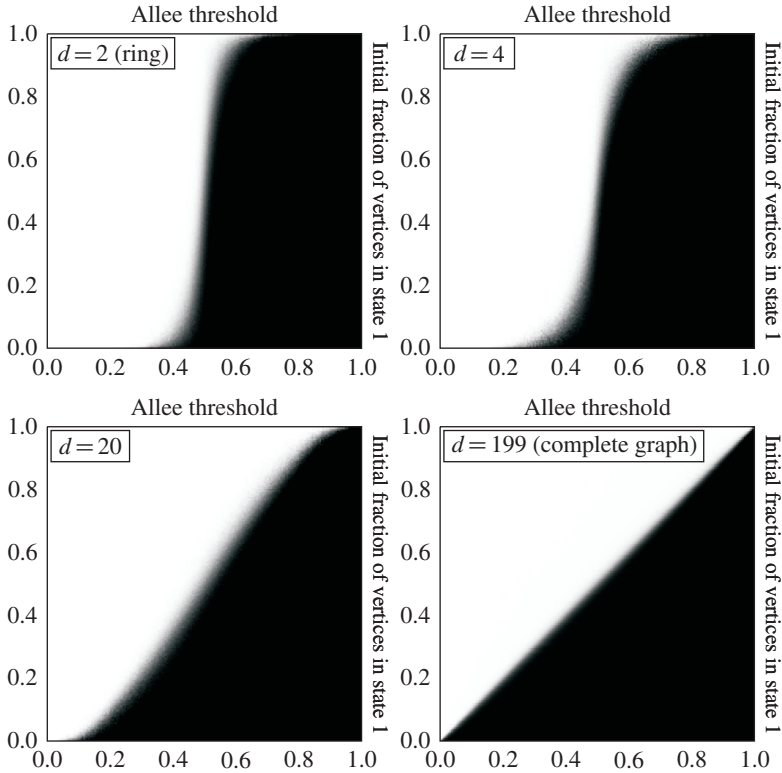
FIGURE 1: Simulation results for the process on the torus $\mathbb{Z}/200\mathbb{Z}$ in which each vertex is connected to its d nearest neighbors. Each picture shows the density of vertices below/above the Allee threshold after a large number of updates as a function of the Allee threshold (200 values) and the initial fraction of vertices in state 1 (200 values). The color for each of the 200×200 parameter values is computed from the average of 100 independent realizations, and the color code is black for all below the Allee threshold and white for all above. In all the simulations, $\mu = 0.2$.

on four regular graphs with different degrees. In contrast, the top-left-hand picture suggests that the critical value for the Allee threshold is again equal to $\frac{1}{2}$ for the process on the ring in the limit as the size of the system tends to ∞. This indicates that the critical value for the Allee threshold strongly depends on the initial configuration in the presence of fast long-range dispersal, whereas the initial configuration is essentially unimportant in the presence of slow short-range dispersal. The main objective of this paper is to explain at least qualitatively this difference between the process on the complete graph, which models fast long-range dispersal, and the process on the ring, which models slow short-range dispersal.

2. Main results

We first assume that the process starts from the product measure in which each vertex is in state 1 with probability ρ and in state 0 otherwise. Following the terminology of [6], we call *global extinction* the event that the process converges to the 'all 0' configuration and *global expansion* the event that it converges to the 'all 1' configuration. Note that, on finite graphs, the process converges to one of these two absorbing states. In particular, the probability of global

extinction and the probability of global expansion sum to 1, thus showing that the long-term behavior is completely characterized by the probability of global expansion

$$p_G(\theta, \mu, \rho) := \mathbb{P}(\eta_t \equiv 1 \text{ for some } t),$$

which depends on the network of interactions, the Allee threshold, the migration factor, and the initial density of occupied patches. The simulation results of Figure 1 suggest some monotonicity of the probability of expansion with respect to the Allee threshold and the initial density of occupied patches, as well as a certain symmetry between the probability of expansion and the probability of extinction. These results follow directly from standard coupling arguments for interacting particle systems that we briefly describe without detailed proof.

Monotonicity with respect to θ. Two processes on the same graph and starting from the same initial configuration but with different Allee thresholds can be coupled in such a way that the process with the smaller Allee threshold dominates the other process, which implies that the probability of global expansion $p_G(\theta, \mu, \rho)$ is nonincreasing with respect to θ.

Monotonicity with respect to ρ. Two processes on the same graph with the same Allee threshold and with the same migration factor can be coupled in such a way that if one process dominates the other at time 0 then the domination remains true at all times. This implies that the probability of global expansion $p_G(\theta, \mu, \rho)$ is nondecreasing with respect to ρ.

Symmetry. The process with Allee threshold θ can be coupled with the process on the same graph and with the same migration factor but with Allee threshold $1 - \theta$ in such a way that if at any vertex the initial population density for one process equals 1 minus the initial population density for the other process then this remains true at all times. This implies that

$$p_G(\theta, \mu, \rho) = 1 - p_G(1 - \theta, \mu, 1 - \rho),$$

and explains the symmetry in the four simulation pictures of Figure 1.

The process starting with a single occupied patch. We now return to the main objective of this paper, which is to understand the effect of the geometry of the network on the invadability of species subject to a strong Allee effect. This aspect is mathematically more difficult to understand because two processes on different graphs cannot be coupled in such a way that one process dominates the other. Our analysis focuses on the extreme cases of the ring and the complete graph corresponding to the top-left- and bottom-right-hand pictures of Figure 1. Let

$$p_N^+(\theta, \mu, \rho) := \text{the probability of global expansion for the process}$$
$$\text{on the complete graph with } N \text{ vertices,} \tag{2a}$$

$$p_N^-(\theta, \mu, \rho) := \text{the probability of global expansion for the process}$$
$$\text{on the ring with } N \text{ vertices,} \tag{2b}$$

where the '+' and '−' superscripts respectively allude to the facts that the complete graph is the connected regular graph with the largest degree, while the ring is the connected regular graph with the smallest degree. Recall that the simulation results of Figure 1 suggest that

$$\lim_{N \to \infty} p_N^+(\theta, \mu, \rho) = \begin{cases} 0 & \text{when } \theta > \rho, \\ 1 & \text{when } \theta < \rho, \end{cases} \tag{3}$$

whereas, for the process on the ring starting with $\rho \in (0, 1)$,

$$\lim_{N \to \infty} p_N^-(\theta, \mu, \rho) = \begin{cases} 0 & \text{when } \theta > \frac{1}{2}, \\ 1 & \text{when } \theta < \frac{1}{2}. \end{cases} \tag{4}$$

Following Kang and Lanchier [6], whose main objective was to understand whether an alien species established in one patch can expand in space, we assume from now on that the process starts with a single patch in state 1 and all the other patches in state 0. In particular, we drop the parameter ρ in the probabilities given in (2). The following two theorems give qualitative differences between the system on the complete graph and the system on the ring starting with a single vertex in state 1, which supports conjectures (3) and (4). More precisely, our first theorem indicates that, even when the Allee threshold is very small, the system on the complete graph is driven toward global extinction with high probability when the number of vertices is large.

Theorem 1. *Assume that $\theta, \mu > 0$. Then $p_N^+(\theta, \mu) \to 0$ as $N \to \infty$.*

In contrast, when the Allee threshold is small enough, the system on the ring expands globally with a positive probability that does not depend on the number of vertices.

Theorem 2. *Assume that $\theta < \mu^2 (1 - \mu)^{1140}$. Then $\inf_N p_N^-(\theta, \mu) > 0$.*

The mysterious assumption in the previous theorem follows from a series of bounds of certain probabilities that are estimated based on geometric arguments and are not optimal. Some of these estimates appear in our proof and the others in the calculation of an upper bound for the critical value of one-dependent oriented site percolation in, e.g. [3]. Even though the assumption of the theorem is far from being optimal, it gives at least an explicit lower bound for the critical value of the Allee threshold for the process on the ring. More importantly, the combination of both theorems shows the following qualitative difference: for some values of the Allee threshold, the probability of global expansion is bounded from below for the process on the ring but vanishes to 0 for the process on the complete graph as the number of vertices increases. This supports at least qualitatively the contrast between (3) and (4). From an ecological perspective, this indicates that, upon arrival of an alien species to a new area, though dispersal is necessary for its expansion, the best strategy is to first slowly disperse to nearby patches, and then progressively increase the strength of its dispersal as the fraction of patches at the carrying capacity increases. Finally, we point out that, since our model assumes that mixing events occur at rate 1 along each edge, increasing the degree of the graph increases not only the range but also the speed of the interactions. An alternative modeling approach is to assume that mixing events occur along each edge at a rate equal to the reciprocal of the degree in order to have local and mixing events occurring at the same time scale. In this case, increasing the degree increases the range but not the speed of the interactions, and it can be proved that Theorem 2 still holds, at least qualitatively, using the exact same arguments. Our proof of Theorem 1, however, does not extend to this model and we conjecture that, for the process on the complete graph, the probability of global expansion is bounded from below uniformly in the number of patches, just as for the process on the ring. In particular, we believe that the qualitative difference that appears in the theorems is due to the speed of the interactions rather than the range of the interactions.

3. Preliminary results

In this section we give some definitions and simple results that will be used repeatedly in the proof of both theorems. Throughout this paper, we think of process (1) as being constructed from Harris' graphical representation [5]. Each edge of the graph is equipped with a Poisson process with intensity 1 while each vertex is equipped with a Poisson process with intensity 1

and a sequence of independent Bernoulli random variables with parameter $\frac{1}{2}$. We write

- $T_n(x, y) := n$th arrival time of the Poisson process attached to edge $(x, y) \in E$,

- $U_n(x) := n$th arrival time of the Poisson process attached to vertex $x \in V$,

- $B_n(x) := n$th member of the Bernoulli sequence attached to vertex $x \in V$.

All these Poisson processes and Bernoulli random variables are independent and together form a percolation structure from which the multipatch model (1) can be constructed.

- *Mixing events.* At time $t := T_n(x, y)$, we draw a double arrow along the corresponding edge to indicate the occurrence of the following mixing event:

$$\eta_t(x) = (\sigma_{x,y}\eta_{t-})(x) \quad \text{and} \quad \eta_t(y) = (\sigma_{x,y}\eta_{t-})(y).$$

- *Local events.* At time $t := U_n(x)$, we put a dot at vertex x to indicate that

$$\eta_t(x) = \mathbf{1}\{\eta_{t-}(x) > \theta\} + B_n(x)\,\mathbf{1}\{\eta_{t-}(x) = \theta\}.$$

In the proof of both theorems, we first study the process (ξ_t) that includes mixing events but excludes local events, whose dynamics are therefore described by the Markov generator

$$L_\xi f(\xi) = \sum_{(x,y)\in E} [f(\sigma_{x,y}\xi) - f(\xi)]. \tag{5}$$

Note that this process can be constructed graphically as previously by using only the Poisson processes attached to the edges of the graph. Note also that, since the state at each vertex is a convex combination of the states of the vertices at earlier times,

$$\xi_s(x) > \theta \quad \text{for all } x \in V \quad \Longrightarrow \quad \xi_t(x) > \theta \quad \text{for all } (x, t) \in V \times (s, \infty),$$

with the analogous implication obtained by flipping the inequalities. Following the terminology introduced in [6], we call the sets

$$\Omega_+ := \{\xi : V \to [0, 1] \text{ such that } \xi(x) > \theta \text{ for all } x \in V\},$$
$$\Omega_- := \{\xi : V \to [0, 1] \text{ such that } \xi(x) < \theta \text{ for all } x \in V\}$$

the *upper/lower configurations*, respectively, and observe that the previous implication means that, once process (5) hits the set of upper configurations, it stays in this set forever. By the definition of the Allee threshold, the same holds for the original process (1) from which we deduce the following lemma.

Lemma 1. *For process (1) on a finite graph,*

$$\eta_t \equiv 1 \quad \text{for some } t \quad \Longleftrightarrow \quad \eta_s \in \Omega_+ \quad \text{for some } s,$$
$$\eta_t \equiv 0 \quad \text{for some } t \quad \Longleftrightarrow \quad \eta_s \in \Omega_- \quad \text{for some } s.$$

Lemma 1 is one of the keys to proving both theorems. According to the lemma, it suffices to prove that, with the appropriate probability, the process on the complete graph hits a lower configuration whereas the process on the ring hits an upper configuration.

4. The process on the complete graph

This section is devoted to the proof of Theorem 1. As previously mentioned, the first step is to study process (5) that excludes local events but includes mixing events. It is obvious that, when starting with a single vertex in state 1, this process eventually hits a lower configuration, provided the number of vertices is sufficiently large. The key to the proof is to show that the time to hit a lower configuration can be made arbitrarily small, which relies on large deviation estimates for several random variables that we now define. We call a mixing event that involves two occupied vertices a *collision*, i.e. a collision occurs at time t whenever

$$t = T_n(x, y) \quad \text{and} \quad \min(\xi_{t-}(x), \xi_{t-}(y)) \neq 0.$$

We define the following three random variables.

- Time to dispersion: $\tau_D := \inf\{t : \xi_t \in \Omega_-\}$.

- Time to collision: $\tau_C :=$ first time a collision occurs.

- Number of occupied patches: $|\xi_t| := \text{card}\{x \in \mathbb{Z}/N\mathbb{Z} : \xi_t(x) \neq 0\}$.

We will prove that, with a probability close to 1 when N is large,

$$\tau_D \leq T_N, \qquad \tau_C > T_N, \quad \text{and} \quad |\xi_{T_N}| \leq 4^{NT_N}, \tag{6}$$

for some T_N that tends to 0 as $N \to \infty$. The probability of the events in (6) will be estimated backwards by conditioning, i.e. the probability involving the time to collision is obtained by conditioning on the number of occupied patches, while the probability involving the time to dispersion is obtained by conditioning on the time to collision. To complete the proof, we will return to process (1) and use our estimates for the probability of the first and last events in (6) to prove that the probability that a local event occurs in any of the occupied patches before the time to dispersion tends to 0 as the number of vertices $N \to \infty$.

Mapping to a dynamic graph. To estimate the probability of the events in (6), we first define a mapping to visualize the evolution of process (5) through a dynamic graph, i.e. a continuous-time Markov chain whose state at time t is a random oriented graph

$$H_t := (V_t, E_t), \quad \text{where} \quad V_t \subset V \times \mathbb{N}.$$

The dynamic graph is coupled with process (5) and defined as follows.

- The graph H_0 has only one vertex, namely $(x_0, 0) \in V \times \mathbb{N}$, where x_0 is the single vertex in state 1 initially, and no (oriented) edge.

- We call (x, i) a *leaf* at time t whenever $(x, i) \in V_t$ and $(x, i + 1) \notin V_t$.

- Assume that $(x, i) \in V_{t-}$ is a leaf and that $t = T_n(x, x')$ for some $n \geq 1$. Then we define the new vertex set and the new edge set as

$$
\begin{aligned}
V_t &:= V_{t-} \cup \{(x, i + 1), (x', i + 1)\}, \\
E_t &:= E_{t-} \cup \{(x, i) \to (x, i + 1), (x, i) \to (x', i + 1)\}.
\end{aligned}
\tag{7}
$$

The left-hand diagram of Figure 2 gives a realization of the process with only mixing events, where x_j denotes the jth vertex that becomes occupied and s_j denotes the time at which this occurs, while the right-hand diagram of the figure gives a picture of the corresponding graph. Note that a vertex has a positive density at time t if and only if it is the first coordinate of a leaf.
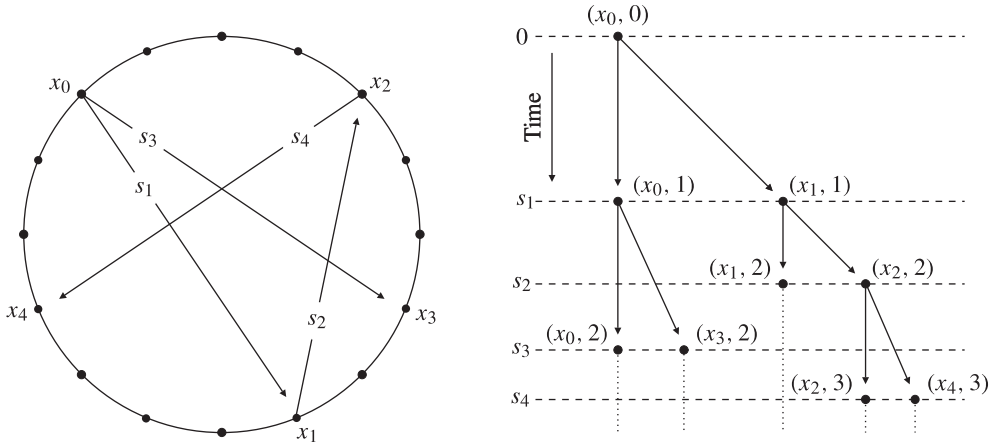
FIGURE 2: Pictures related to the proof of Theorem 1.

It follows that

$$|\xi_t| = \text{ number of leaves in } H_t \tag{8}$$

and, using in addition the evolution rule (7), that a collision event results in two pairs of oriented edges, each pointing to the same two leaves. In particular,

$$H_t \text{ is an oriented binary tree with root } (x_0, 0) \text{ if and only if } t < \tau_C. \tag{9}$$

From properties (7) and (9), it also follows that

$$(x, i) \text{ is a leaf at time } t < \tau_C \implies \xi_t(x) \le (1 - \mu)^i. \tag{10}$$

In fact, if (x, i) is a leaf at time $t < \tau_C$ then there exists a unique oriented path from the root to this leaf and the density at vertex x can be computed explicitly by looking at the number of vertical edges in this path, but this property is not needed in the proof of the theorem.

Time to dispersion. The next step is to use (8)–(10) and the dynamic graph representation of the process to estimate the probability of the three events in (6). Under the assumptions of the theorem, there exists n such that $(1 - \mu)^n < \theta$. We then define

$$T_N := \frac{n \ln(\ln N)}{N} \quad \text{and} \quad K_N := 4^{n \ln(\ln N)} = 4^{NT_N}.$$

The following three lemmas give estimates of the probability of the last event, the second event, and the first event in (6), respectively, for the deterministic time T_N defined above.

Lemma 2. *There exists $a > 0$ such that $\mathbb{P}(|\xi_{T_N}| > K_N) \le (\ln N)^{-a}$ for all large N.*

Proof. The number of leaves is maximal when there is no collision, in which case the number of leaves jumps from i to $i + 1$ at rate iN. This together with (8) implies that

$$\mathbb{E}|\xi_{T_N}| = \mathbb{E}(\text{number of leaves in } H_{T_N}) \le 2^{NT_N} = \sqrt{K_N}.$$

In particular, large deviation estimates for the Poisson distribution give

$$\mathbb{P}(|\xi_{T_N}| > K_N) = \mathbb{P}(|\xi_{T_N}| > 4^{NT_N}) \le \exp(-a \ln(\ln N)) = (\ln N)^{-a}$$

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

Lemma 3. *For all large N, $\mathbb{P}(\tau_C \le T_N) \le 2(\ln N)^{-a}$.*

Proof. Given that the graph H_t has i leaves, the probability of a collision at the next update of the system is equal to i/N. In particular, the conditional probability of a collision before T_N, given that the number of leaves at that time is smaller than K_N, is

$$\mathbb{P}(\tau_C \leq T_N \mid |\xi_{T_N}| \leq K_N) \leq \sum_{i=1}^{K_N} \frac{i}{N} \leq \frac{K_N(K_N+1)}{2N} \leq \frac{\exp(4n \ln(\ln N))}{2N} = \frac{(\ln N)^{4n}}{2N}.$$

This, together with Lemma 2, implies that

$$\begin{aligned} \mathbb{P}(\tau_C \leq T_N) &\leq \mathbb{P}(\tau_C \leq T_N \mid |\xi_{T_N}| \leq K_N) + \mathbb{P}(|\xi_{T_N}| > K_N) \\ &\leq \frac{(\ln N)^{4n}}{2N} + (\ln N)^{-a} \\ &\leq 2(\ln N)^{-a} \end{aligned}$$

for all sufficiently large N.

Lemma 4. *Let $a > 0$ be as in Lemmas 2 and 3. Then, for all large N,*

$$\mathbb{P}(\tau_D > T_N) \leq 2^{n+1}(\ln N)^{-1} + 2(\ln N)^{-a}.$$

Proof. Motivated by (10), we introduce the stopping times

$$\sigma_j := \inf\{t \colon \text{for each leaf } (x, i) \in V_t, \text{ we have } i \geq j\}.$$

The first step is to prove that $\sigma_n \leq T_N$ with probability arbitrarily close to 1 when the number of vertices is large. Note that, according to the evolution rules (7), we have

$$\operatorname{card}\{x \in V \colon (x, i) \in V_t\} \leq 2^i \quad \text{for all } t \geq 0. \tag{11}$$

Moreover, since each vertex has degree $N - 1$ and is therefore involved in a mixing event at the arrival times of a Poisson process with intensity $N - 1$, we have

$$\mathbb{P}(\inf\{t \colon (x, i+1) \in V_t\} - \inf\{t \colon (x, i) \in V_t\} > T) = \exp(-(N-1)T) \tag{12}$$

for all $x \in V$ such that $(x, i) \in V_t$ at some time t, i.e. the amount of time a vertex in the dynamic graph is a leaf is exponential with parameter $N - 1$. From (11) and (12), we deduce that the temporal increment $\sigma_{j+1} - \sigma_j$ required to grow one more generation in the dynamic graph is stochastically smaller than the maximum of 2^j independent exponential random variables with the same parameter $N - 1$. In particular, having a collection $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$ of independent exponential random variables with parameter $N - 1$, we deduce that

$$\begin{aligned} \mathbb{P}(\sigma_n > T_N) &\leq \sum_{j=0}^{n-1} \mathbb{P}\left(\sigma_{j+1} - \sigma_j > \frac{T_N}{n}\right) \\ &\leq \sum_{j=0}^{n-1} \mathbb{P}\left(\max\{\varepsilon_i \colon i \leq 2^j\} > \frac{T_N}{n}\right) \\ &\leq \sum_{j=0}^{n-1} \mathbb{P}\left(\varepsilon_i > \frac{T_N}{n} \text{ for some } i \leq 2^j\right) \\ &\leq \sum_{j=0}^{n-1} 2^j \exp\left(-\frac{(N-1)\ln(\ln N)}{N}\right) \\ &\leq 2^{n+1}(\ln N)^{-1} \quad \text{for all large } N. \end{aligned} \tag{13}$$

In other respects, using (10) and recalling the definition of n,

$$
\sigma_n \le T_N < \tau_C \quad \Longrightarrow \quad \xi_{T_N}(x) \le (1-\mu)^n < \theta \quad \text{for each leaf } (x, i) \in V_{T_N}
$$
$$
\Longrightarrow \quad \tau_D \le T_N. \tag{14}
$$

Combining (13) and (14), and using Lemma 3, we conclude that

$$
\mathbb{P}(\tau_D > T_N) \le \mathbb{P}(\sigma_n > T_N) + \mathbb{P}(\tau_C \le T_N) \le 2^{n+1}(\ln N)^{-1} + 2(\ln N)^{-a},
$$

which completes the proof.

To conclude the proof of Theorem 1, we return to process (1). The proof is based on the simple observation that processes (1) and (5) are equal as long as no local event occurs in any of the vertices not in state 0. The bound on the number of occupied patches and the bound on the time to dispersion respectively given by Lemmas 2 and 4 show that, with probability close to 1 when N is large, no local event occurs in any of the occupied patches before the time to dispersion (for the process that includes local events), so the result follows from Lemma 1. This argument is made rigorous in the following lemma.

Lemma 5. *For all sufficiently large N, we have*

$$
p_N^+(\theta, \mu) \le 2^{n+1}(\ln N)^{-1} + 3(\ln N)^{-a} + \frac{n(\ln N)^{2n+1}}{N}. \tag{15}
$$

Proof. The two processes (1) and (5) being constructed from the same graphical representation are equal as long as no local event occurs in any of the vertices not in state 0. Since local events occur at each vertex at rate 1, having a collection $\zeta_1, \zeta_2, \zeta_3, \ldots$ of independent exponential random variables with parameter 1, we deduce that

$$
\begin{aligned}
\mathbb{P}(\eta_t \not\equiv \xi_t \text{ for some } t \le T_N \mid |\xi_{T_N}| \le K_N) &\le \mathbb{P}(\min\{\zeta_i : i \le K_N\} < T_N) \\
&= 1 - \mathbb{P}(\min\{\zeta_i : i \le K_N\} \ge T_N) \\
&\le 1 - \exp(-K_N T_N) \\
&\le 1 - \exp\left(-\frac{n(\ln N)^{2n} \ln(\ln N)}{N}\right) \\
&\le \frac{n(\ln N)^{2n+1}}{N}
\end{aligned} \tag{16}
$$

for all large N. From Lemmas 2 and 4, and (16), we obtain

$$
\begin{aligned}
\mathbb{P}(\eta_{T_N} \notin \Omega_-) &\le \mathbb{P}(\xi_{T_N} \notin \Omega_-) + \mathbb{P}(\eta_t \not\equiv \xi_t \text{ for some } t \le T_N) \\
&\le \mathbb{P}(\tau_D > T_N) + \mathbb{P}(\eta_t \not\equiv \xi_t \text{ for some } t \le T_N \mid |\xi_{T_N}| \le K_N) \\
&\quad + \mathbb{P}(|\xi_{T_N}| > K_N) \\
&\le 2^{n+1}(\ln N)^{-1} + 3(\ln N)^{-a} + \frac{n(\ln N)^{2n+1}}{N}
\end{aligned}
$$

for all large N. Since according to Lemma 1 we have

$$
p_N^+(\theta, \mu) = \mathbb{P}(\eta_t \in \Omega_+ \text{ for some } t) = \mathbb{P}(\eta_t \notin \Omega_- \text{ for all } t) \le \mathbb{P}(\eta_{T_N} \notin \Omega_-),
$$

the proof is complete.

The theorem directly follows from Lemma 5 by observing that the right-hand side of (15) tends to 0 as the number of vertices N goes to ∞.

5. The process on the ring

This section is devoted to the proof of Theorem 2. To understand the process on the ring, the first step is to study its counterpart on the infinite one-dimensional lattice using a so-called block construction. The idea is to couple a certain collection of good events related to the infinite system with the set of open sites of a one-dependent oriented site percolation process on

$$\mathcal{H} := \{(x, n) \in \mathbb{Z} \times \mathbb{Z}_+ : x + n \text{ is even}\}.$$

For a precise definition and a review of oriented site percolation in two dimensions, we refer the reader to [3]. This coupling together with results from [4] implies that, starting with a single occupied patch, there exists, with positive probability, a linearly expanding region that contains a positive density of patches above the Allee threshold. The second key step is to prove that in fact all the patches in this space–time region are above the Allee threshold, from which it follows that, with the same positive probability, the process on the finite ring starting with a single occupied patch hits an upper configuration before it hits a lower configuration. To prove linear expansion in space of the set of patches that exceed the Allee threshold, we observe that, under the assumptions of the theorem, there exists a constant a fixed from now on such that

$$a < (1 - \mu)^{4T} \quad \text{and} \quad a\mu^2(1 - \mu)^{8T} > \theta, \quad \text{where} \quad T := 95.$$

To define our collection of good events, we also introduce random variables that keep track of the number of mixing events and local events in certain space–time regions of the graphical representation. More precisely, we introduce the number of mixing events

$$X_j := \text{card}\{n : T_n(j, j + 1) \in (0, T)\} \quad \text{for } j = -1, -2,$$
$$X_j := \text{card}\{n : T_n(j - 1, j) \in (0, T)\} \quad \text{for } j = +1, +2,$$
$$Y_j := \text{card}\{n : T_n(j - 1, j) \in (T, 2T)\}$$
$$+ \text{card}\{n : T_n(j, j + 1) \in (T, 2T)\} \quad \text{for } j = -1, +1,$$

as well as the number of local events

$$Z_j := \text{card}\{n : U_n(j) \in (T, 2T)\} \quad \text{for } j = -1, +1.$$

From these random variables, we define the good event

$$\Omega := \{\min(X_{-1}, X_1, Z_{-1}, Z_1) \neq 0\} \cap \{\max(X_{-2}, X_2) \leq 2T\} \cap \{\max(Y_{-1}, Y_1) \leq 4T\},$$

as depicted in Figure 3. For every $(x, n) \in \mathcal{H}$, we define the good event $\Omega(x, n)$ similarly but from the graphical representation of the process in the space–time region

$$R(x, n) := (x - 2, x + 2) \times (2nT, 2nT + 2T).$$

The motivation for introducing these events is that, conditioned on $\Omega(x, n)$, if the population density at patch x at time $2nT$ exceeds a then the same holds for the two adjacent patches $2T$ units of time later. By translation invariance of the evolution rules of the process in space and time, it suffices to prove the result for $x = n = 0$, which is done in the following lemma.

Lemma 6. *We have* $\{\eta_0(0) > a\} \cap \Omega(0, 0) \subset \{\min(\eta_{2T}(-1), \eta_{2T}(1)) > a\}$.

Proof. The first step is to prove that the event on the left-hand side is included in the event that the population density at -1 and 1 exceed the Allee threshold at time $2T$ for process (5)
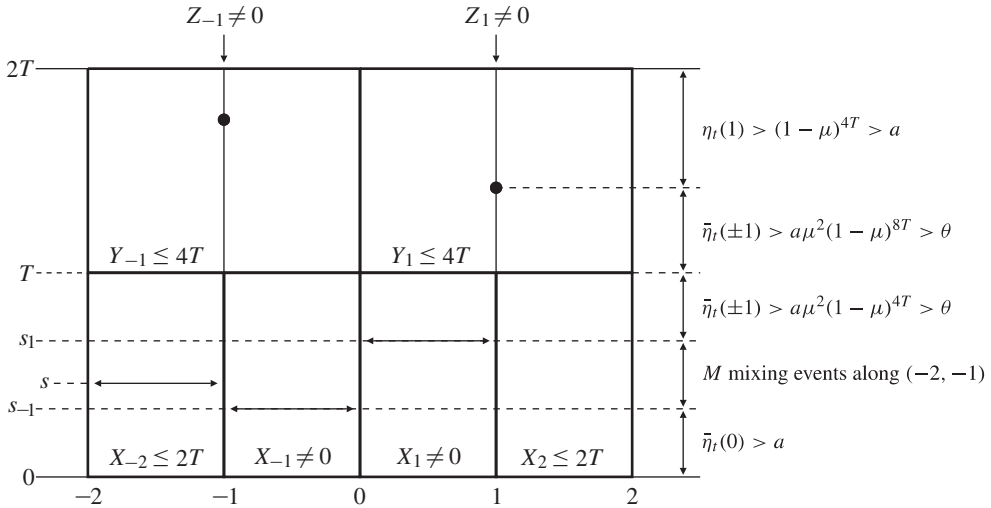
FIGURE 3: Picture of the good event $\Omega = \Omega(0,0)$.

that includes mixing events but excludes local events. Assume that

$$s_{-1} := T_1(-1, 0) < T_1(0, 1) =: s_1 < T.$$

Note that the last inequality $s_1 < T$ follows from $X_1 \neq 0$. To study the process up until time s_1, we introduce the number of mixing events

$$M := \mathrm{card}\{n : T_n(-2, -1) \in (s_{-1}, s_1)\},$$

and in the case $M \neq 0$, the time of the first mixing event

$$s := \inf(\{T_n(-2, -1) : n \geq 1\} \cap (s_{-1}, s_1)).$$

Since the population densities at patch -1 and at patch 0 between time s_{-1} and time s are convex combinations of their counterpart at time s_{-1}, we have

$$\min(\xi_t(-1), \xi_t(0)) \geq \min(\xi_{s_{-1}}(-1), \xi_{s_{-1}}(0)) > \min(a\mu, a(1-\mu)) = a\mu$$

for all $t \in (s_{-1}, s)$. In particular,

$$\min(\xi_s(-1), \xi_s(0)) > a\mu(1-\mu).$$

Since $M \leq X_{-2} \leq 2T$, we deduce from a simple induction that

$$\min(\xi_t(-1), \xi_t(0)) > a\mu(1-\mu)^M \geq a\mu(1-\mu)^{2T} > \theta \tag{17}$$

for all $t \in (s_{-1}, s_1)$, and, by the definition of s_1,

$$\xi_{s_1}(1) = \mu\xi_{s_1-}(0) + (1-\mu)\xi_{s_1-}(1) \geq \mu\xi_{s_1-}(0) > a\mu^2(1-\mu)^M.$$

In particular, again using our inductive reasoning and the fact that

$$\mathrm{card}\{n : T_n(-2, -1) \in (s_1, T)\} + \mathrm{card}\{n : T_n(1, 2) \in (s_1, T)\} \leq X_{-2} + X_2 - M$$
$$\leq 4T - M,$$

we deduce that

$$
\begin{aligned}
\min(\xi_t(-1), \xi_t(0), \xi_t(1)) &\geq (1-\mu)^{4T-M} \min(\xi_{s_1}(-1), \xi_{s_1}(0), \xi_{s_1}(1)) \\
&> (1-\mu)^{4T-M} a\mu^2 (1-\mu)^M \\
&= a\mu^2 (1-\mu)^{4T} \\
&> \theta
\end{aligned}
\tag{18}
$$

for all $t \in (s_1, T)$. Finally, since $Y_{-1} \leq 4T$ and $Y_1 \leq 4T$,

$$
\begin{aligned}
\min(\xi_t(-1), \xi_t(1)) &\geq (1-\mu)^{4T} \min(\xi_T(-1), \xi_T(1)) \\
&> (1-\mu)^{4T} a\mu^2 (1-\mu)^{4T} \\
&= a\mu^2 (1-\mu)^{8T} \\
&> \theta
\end{aligned}
\tag{19}
$$

for all $t \in (T, 2T)$. Returning to the system with local events, since

$$
\xi_t(0) > \theta \quad \text{for all } t \in (0, T) \quad \text{and} \quad \xi_t(\pm 1) > \theta \quad \text{for all } t \in (s_{\pm 1}, 2T)
$$

according to (17)–(19), these inequalities remain true for the original process (1). In particular, the populations at patches -1 and 1 exceed θ between times T and $2T$; therefore,

$$
\eta_t(-1) = 1 \quad \text{for some } t \in (T, 2T) \quad \text{and} \quad \eta_t(1) = 1 \quad \text{for some } t \in (T, 2T)
$$

since $Z_{-1} \neq 0$ and $Z_1 \neq 0$. Again using $Y_{-1}, Y_1 \leq 4T$, we conclude that

$$
\eta_{2T}(-1) \geq (1-\mu)^{4T} > a \quad \text{and} \quad \eta_{2T}(1) \geq (1-\mu)^{4T} > a,
$$

which completes the proof of the lemma.

To deduce from Lemma 6 the existence of a linearly expanding region with a positive density of patches above the Allee threshold, we now prove that the common probability of all our good events exceeds the critical value p_c of one-dependent oriented site percolation.

Lemma 7. *For $T = 95$, we have $\mathbb{P}(\Omega(x, n)) \geq 1 - 3^{-36} > p_c$.*

Proof. The key is simply to observe that

- X_{-2}, X_{-1}, X_1, and X_2 are Poisson random variables with parameter T,
- Y_{-1} and Y_1 are Poisson random variables with parameter $2T$, and
- Z_{-1} and Z_1 are Poisson random variables with parameter T.

Using, in addition, the fact that these random variables are independent, we deduce that

$$
\begin{aligned}
\mathbb{P}(\Omega(x, n)) = \mathbb{P}(\Omega) &\geq 1 - 4\mathbb{P}(X_1 = 0) - 2\mathbb{P}(X_2 > 2T) - 2\mathbb{P}(Y_1 > 4T) \\
&= 1 - 4e^{-T} - 2 \sum_{n>2T} \frac{T^n}{n!} e^{-T} - 2 \sum_{n>4T} \frac{(2T)^n}{n!} e^{-2T} \\
&> 1 - 3^{-36}
\end{aligned}
$$

when $T = 95$. The second inequality $p_c < 1 - 3^{-36}$ in the statement follows from the contour argument described in, e.g. [3, Section 10]. This completes the proof of the lemma.

Lemmas 6 and 7 and the fact that

$$R(x, n) \cap R(x', n') = \varnothing \quad \text{whenever } |x - x'| > 1 \text{ or } n \neq n'$$

are the assumptions of Theorem 4.3 of [4] with $M = 1$ and $\gamma = 3^{-36}$, from which it follows that, for the infinite system starting with a single occupied patch at the origin,

$$\bar{W}_n := \{x \in \mathbb{Z} : (x, n) \in \mathcal{H} \text{ and } \eta_{2nT}(x) > a\}$$

dominates stochastically the set of wet sites W_n at level n of a one-dependent oriented site percolation process with parameter $1 - \gamma$ and initial condition $W_0 = \bar{W}_0$. Since $1 - \gamma > p_c$, we deduce that, with positive probability at least equal to the percolation probability, the set of patches that exceed the constant a expands linearly. This proves only persistence of the metapopulation of the infinite lattice, which is not sufficient to deduce global expansion of the system on the ring. The last step is to show that all patches in the expanding region exceed the Allee threshold. More precisely, on the event that percolation occurs, i.e. $W_n \neq \varnothing$ for all n, we have

$$\lim_{n \to \infty} l_n := \lim_{n \to \infty} \min W_n = -\infty \quad \text{and} \quad \lim_{n \to \infty} r_n := \lim_{n \to \infty} \max W_n = +\infty, \qquad (20)$$

and thinking of the infinite system as being coupled with one-dependent oriented site percolation in such a way that $W_n \subset \bar{W}_n$ for all n, we have the following lemma.

Lemma 8. *Assume that $W_n \neq \varnothing$ for all n. Then, for all $x \in \mathbb{Z}$,*

$$\eta_t(x) > \theta \quad \textit{for all sufficiently large times } t.$$

Proof. Recall from the proof of Theorem 4.3 of [4] that the processes are coupled in such a way that the set of open sites for the percolation process is included in the set of good sites, where site (x, n) is said to be *good* whenever the good event $\Omega(x, n)$ occurs. This, together with the definition of the right edge, implies that there is a good path from site $(0, 0)$ to site (r_n, n), i.e. a sequence of integers $x_0 = 0, x_1, \ldots, x_n = r_n$ such that

$$(x_m, m) \text{ is good for } m = 0, 1, \ldots, n \quad \text{and} \quad |x_m - x_{m-1}| = 1 \quad \text{for } m = 1, 2, \ldots, n.$$

Since $\eta_0(0) > a$, it follows from (17)–(19) in the proof of Lemma 6 that, for $m = 0, 1, \ldots, n$,

$$\begin{aligned}
\eta_t(x_m) > \theta \quad &\text{for all } t \in [2mT, 2mT + T), \\
\eta_t(x_m + 1) > \theta \quad &\text{for all } t \in [2mT + T, 2mT + 2T).
\end{aligned} \qquad (21)$$

Similarly, there is a path $(x_0, 0) \to (x_{-1}, 1) \to \cdots \to (x_{-n}, n) = (l_n, n)$ such that

$$\begin{aligned}
\eta_t(x_{-m}) > \theta \quad &\text{for all } t \in [2mT, 2mT + T), \\
\eta_t(x_{-m} - 1) > \theta \quad &\text{for all } t \in [2mT + T, 2mT + 2T),
\end{aligned} \qquad (22)$$

for $m = 0, 1, \ldots, n$, and we may assume that $x_{-m} \leq x_m$ for all m. We claim that all patches in the space–time region delimited by (21) and (22) are above the Allee threshold θ, i.e.

$$\begin{aligned}
\eta_t(x) > \theta \quad &\text{for all } (x, t) \in [x_{-m}, x_m] \times [2mT, 2mT + T), \\
\eta_t(x) > \theta \quad &\text{for all } (x, t) \in [x_{-m} - 1, x_m + 1] \times [2mT + T, 2mT + 2T),
\end{aligned} \qquad (23)$$

which we prove by induction. Assume that (23) holds for some $m < n$. The fact that this again holds at time $t = 2mT + 2T$ simply follows from the fact that

$$[x_{-m-1}, x_{m+1}] \subset [x_{-m} - 1, x_m + 1] \quad \text{since} \quad x_{-m-1} \geq x_{-m} - 1 \quad \text{and} \quad x_{m+1} \leq x_m + 1.$$

To prove that this holds at later times, we distinguish three types of event.

- Local events cannot violate the first line of (23) since patches above the Allee threshold can experience only local expansions to their carrying capacity.

- Mixing events in $[x_{-m-1}, x_{m+1}]$ cannot violate the first line of (23) since the new states of interacting patches above the Allee threshold are again above the Allee threshold.

- Mixing events along $(x_{-m-1} - 1, x_{-m-1})$ or $(x_{m+1}, x_{m+1} + 1)$ can violate the first line of (23), but this would contradict either (21) or (22).

This proves the first line of (23) at step $m + 1$, while the second line follows from the exact same reasoning. From (23), we deduce that

$$\eta_t(x) > \theta \quad \text{for all } (x, t) \in [l_n, r_n] \times [2mT, 2mT + 2T).$$

In particular, the lemma follows from (20).

Returning to the process on the ring, we deduce from Lemma 8 that, with positive probability at least equal to the percolation probability and starting with a single patch in state 1, the system reaches an upper configuration before it reaches a lower configuration, an event that leads to global expansion according to Lemma 1. This completes the proof of Theorem 2.

Acknowledgement

References

[1] BORRELLO, D. (2012). On the role of Allee effect and mass migration in survival and extinction of a species. *Ann. Appl. Prob.* **22,** 670–701.

[2] COURCHAMP, F., BEREC, L. AND GASCOIGNE, J. (2009). *Allee Effects in Ecology and Conservation*. Oxford University Press.

[3] DURRETT, R. (1984). Oriented percolation in two dimensions. *Ann. Prob.* **12,** 999–1040.

[4] 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.

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

[6] KANG, Y. AND LANCHIER, N. (2011). Expansion or extinction: deterministic and stochastic two-patch models with Allee effects. *J. Math. Biol.* **62,** 925–973.

[7] SCHINAZI, R. B. (2005). Mass extinctions: an alternative to the Allee effect. *Ann. Appl. Prob.* **15,** 984–991.

[8] SCHINAZI, R. B. (2008). On the role of spatial aggregation in the extinction of a species. In *In and Out of Equilibrium. 2* (Progress Prob. **60**), Birkhäuser, Basel, pp. 551–557.