J. Math. Biol. (2011) 62:925–973
DOI 10.1007/s00285-010-0359-3

Mathematical Biology

Expansion or extinction: deterministic and stochastic two-patch models with Allee effects

Yun Kang · Nicolas Lanchier

Received: 8 February 2010 / Revised: 9 July 2010 / Published online: 3 August 2010

Abstract We investigate the impact of Allee effect and dispersal on the long-term evolution of a population in a patchy environment. Our main focus is on whether a population already established in one patch either successfully invades an adjacent empty patch or undergoes a global extinction. Our study is based on the combination of analytical and numerical results for both a deterministic two-patch model and a stochastic counterpart. The deterministic model has either two, three or four attractors. The existence of a regime with exactly three attractors only appears when patches have distinct Allee thresholds. In the presence of weak dispersal, the analysis of the deterministic model shows that a high-density and a low-density populations can coexist at equilibrium in nearby patches, whereas the analysis of the stochastic model indicates that this equilibrium is metastable, thus leading after a large random time to either a global expansion or a global extinction. Up to some critical dispersal, increasing the intensity of the interactions leads to an increase of both the basin of attraction of the global extinction and the basin of attraction of the global expansion. Above this threshold, for both the deterministic and the stochastic models, the patches tend to synchronize as the intensity of the dispersal increases. This results in either a global expansion or a global extinction. For the deterministic model, there are only two attractors, while the stochastic model no longer exhibits a metastable behavior. In the presence of strong dispersal, the limiting behavior is entirely determined by the value of the Allee thresholds as the global population size in the deterministic and the

Y. Kang (✉)
Applied Sciences and Mathematics, Arizona State University, Mesa, AZ 85212, USA
e-mail: yun.kang@asu.edu

N. Lanchier
School of Mathematical and Statistical Sciences, Arizona State University, P.O. Box 871804, Tempe, AZ 85287, USA
e-mail: lanchier@math.asu.edu

stochastic models evolves as dictated by their single-patch counterparts. For all values of the dispersal parameter, Allee effects promote global extinction in terms of an expansion of the basin of attraction of the extinction equilibrium for the deterministic model and an increase of the probability of extinction for the stochastic model.

Keywords Allee effect · Deterministic model · Stochastic model · Extinction · Expansion · Invasion · Bistability · Basin of attraction · Metastability

Mathematics Subject Classification (2000) 92D25 · 34D45 · 60K35

1 Introduction

Biological invasions of alien species are commonly divided into three stages: arrival, establishment, and expansion (Liebhold and Tobin 2008). The precise circumstances of an alien species' arrival, which refers to the transport of an alien species to new areas outside of its native range, are generally not known and are not the purpose of this article. The establishment stage refers to a growth phase of the population density up to some threshold above which it is usually assumed that natural extinction is highly unlikely. However, if during the expansion stage, which refers to the spreading of the alien species to nearby new areas, the population expands in space through dispersal without significantly increasing its size, thus leading to a drop of its density, there might be a risk of extinction for species subject to an Allee effect. The Allee effect refers to a certain process that leads to decreasing net population growth with decreasing density, thus inducing the existence of a so-called Allee threshold below which populations are driven toward extinction (Courchamp et al. 2009). The causes of Allee effect identified by ecologists are numerous. They include failure to locate mates (Hopper and Roush 1993; Berec et al. 2001), inbreeding depression (Lande 1998), failure to satiate predators (Gascoigne and Lipcius 2004), lack of cooperative feeding (Clark and Faeth 1997), etc. Stochasticity, e.g., demographic and/or environmental stochasticity, may also play an important role during the critical time period when an alien species already established in one area starts to spread its population into a new area through dispersal. In this article, we think of the establishment stage as a local expansion of the population in a given geographical location, which involves an increase of population density in this location, while we think of the expansion stage as a global expansion of the population in space into nearby geographical locations regardless of its density. We call a global expansion successful if it leads to the population being established in nearby geographical locations, and unsuccessful if on the contrary the population fails to get established in new locations which may also lead to a global extinction (the population goes extinct in all patches). The main purpose of this article is to study the critical time period when a species already established in a specific geographical location starts to expand in space, and determine whether the expansion stage is successful or not. Both Allee effect and stochasticity are central to better understand why some alien species successfully expand into new geographical areas, and there has been recently a growing recognition

of the importance of these two components in biological invasions (Drake 2004; Leung et al. 2004; Taylor and Hastings 2005; Ackleh et al. 2007; Courchamp et al. 2009). Understanding their role and strength is of critical importance to gain some insight into why some species are more invasive than others, and may suggest some proper biological control strategies to regulate some populations (Liebhold and Tobin 2008).

If an alien species subject to an Allee effect establishes its population in one area, i.e., its population is above the Allee threshold in this area, then the first step of population expansion is to spread to a nearby new area where the population is either absent or at least below the Allee threshold. A natural way to model this situation is to consider a two-patch model with heterogeneous initial conditions such that

1. both patches are coupled by interacting through dispersal, and
2. in the absence of interactions, i.e., when the patches are uncoupled, the initial conditions lead to establishment in one patch and extinction in the other patch.

This approach has been used previously by Adler (1993) and Kang and Armbruster (2010). In this article, we follow this modeling strategy to study the global expansion and global extinction of an alien species subject to an Allee effect during the critical time period between the establishment stage and the expansion stage by employing both a deterministic two-patch model and a stochastic counterpart. The objectives of our study are twofold: the first is to study the consequences of the inclusion of dispersal and Allee effect on the extinction and expansion for both deterministic and stochastic models with heterogeneous initial conditions; the second is to understand the effects of stochasticity by comparing the results based on both models.

There is a copious amount of literature on the invasion and extinction of populations subject to Allee effects (e.g., Dennis 1989, 2002; Veit and Lewis 1996; McCarthy 1997; Shigesada and Kawasaki 1997; Greene and Stamps 2001; Keitt et al. 2001; Fagan et al. 2002; Wang et al. 2002; Liebhold and Bascompte 2003; Schreiber 2003; Zhou et al. 2004; Petrovskii et al. 2005; Taylor and Hastings 2005) which also includes various models in patchy environment (e.g., Amarasekare 1998a,b; Gyllenberg et al. 1999; Ackleh et al. 2007; Kang and Armbruster 2010).

In the deterministic side, Amarasekare (1998a,b) investigated how an interaction between local density dependence, dispersal, and spatial heterogeneity influence population persistence in patchy environments. In particular, she studied how Allee (or Allee-like) effects arise from these patchy models. Gyllenberg et al. (1999) studied a deterministic model of a symmetric two-patch metapopulation to determine conditions that allow the Allee effect to conserve and create spatial heterogeneities in population densities. Rather than exploring the global dynamics of their models, both Amarasekare (1998a,b) and Gyllenberg et al. (1999) studied the influence of an Allee effect on local dynamics, e.g., number of equilibria and local stability. There are few studies regarding the influence of an Allee effect on the extinction versus expansion of populations in patchy environments (e.g., Ackleh et al. 2007; Kang and Armbruster 2010). Kang and Armbruster (2010) studied the influence of an *Allee-like effect* for a discrete-time two-patch model on plant-herbivore interactions where patches are coupled through a dispersal. Their study suggests that for a certain range of dispersal parameters the population of herbivores in both patches drops under the

Allee threshold, thus leading to an extinction of the herbivores in both patches, for the majority of positive initial conditions.

In the stochastic side, the recent work by Ackleh et al. (2007) focuses on a multi-patch population model combining stochasticity and Allee effect. Their numerical simulations show that populations with initial sizes below but near their Allee threshold in each patch can still become established and invasive if stochastic processes affect life history parameters. The closer the population to its Allee threshold, the greater the probability of invasion. A more theoretical approach based on interacting particle systems can be found in Krone (1999). In his model, each site of the infinite integer lattice has to be thought of as a patch which is either empty, occupied by a small colony with a high risk of going extinct, or occupied by a full colony with a longer life span. If successful, a small colony gets established to become a full colony, while empty patches get colonized by a small colony due to invasions from adjacent full colonies, making space explicit.

In this paper, although we model the population dynamics deterministically following the approach of Amarasekare (1998a,b), Gyllenberg et al. (1999) and Ackleh et al. (2007), our stochastic process as well as analytical results for both models are new. For the deterministic model, our focus is on the global dynamics of the system combining dispersal and Allee effects. In particular, we give analytical results on how Allee threshold and dispersal affect the geometry of the basins of attraction of the stable equilibria. The stochastic model is closely related to the deterministic one and consists of a process that has two absorbing states corresponding to global extinction and global expansion, which allows to have a rigorous definition of successful invasion. In particular, our model is designed to study analytically the probability that a fully occupied patch successfully invade a nearby empty patch. To gain insight into the effects of stochasticity on the population dynamics, we will compare in detail the results obtained for both models.

The rest of the article is organized as follows. In Sect. 2, we introduce the deterministic two-patch model with Allee effect coupled by dispersal. Based on the analysis of the invariant sets, we give a complete picture of the global dynamics of the system including the existence of the nontrivial locally stable equilibria and the geometry of their basin of attraction. Numerical solutions of the deterministic model are given to gain some insight into how dispersal and Allee threshold affect the exact basin of attraction of the equilibria. In Sect. 3, we introduce and analyze mathematically the stochastic model focusing on the time to absorption of the process, the existence of metastable states and the probability of a successful invasion when starting from heterogeneous initial conditions. Numerical simulations of the stochastic model have also been performed to better understand these aspects. In Sect. 4, we introduce, along with additional numerical results, more general deterministic and stochastic models including environmental heterogeneities as the Allee thresholds are different in each patch. In Sect. 5, we compare the predictions based on both models, and describe the biological implications of our analytical and numerical results. Finally, Sects. 6 and 7 are devoted to the proofs of the results related to the deterministic and stochastic models, respectively.

2 A simple deterministic two-patch model with Allee effects

The first step in constructing the deterministic two-patch model is to consider single-species dynamics including an Allee affect as potential candidates to describe the evolution in a single patch. The two-patch model is then naturally derived by looking at a two-dimensional system in which both components are coupled through dispersal. The ecological dynamics of a single species' population subject to an Allee effect that can mimic the dynamics in the absence of dispersal is usually described by the model

$$x' = G(x)\,x - H(x) \tag{1}$$

where $x(t)$ denotes the population density at time t. The function G measures the logistic component of population growth, which is given by

$$G(x) = r - ax \tag{2}$$

where r is the per capita intrinsic growth rate and a measures the extra mortality caused by intraspecific competition. In general, the bistability of the differential equation (1) is triggered by combining the negative density-dependence of the logistic growth G with the positive density-dependence of an additional demographic factor represented here by the function H. The decreasing reproduction due to a shortage of mating encountered in low population density and the decreasing mortality due to the weakening predation risk in higher population density are two important examples of such factors (Stephens and Sutherland 1999) which, following Dercole et al. (2002), can be modeled by a Holling type II functional response: $H(x) = cx/(x+d)$. This gives the following population model

$$x' = x(r - ax) - \frac{cx}{x+d} \tag{3}$$

which can be rewritten as

$$\begin{aligned} x' = \frac{ax}{x+d}\left(x - \frac{r - ad - \sqrt{(r+ad)^2 - 4ac}}{2a}\right) \\ \times\left(\frac{r - ad + \sqrt{(r+ad)^2 - 4ac}}{2a} - x\right) \end{aligned} \tag{4}$$

provided $(r+ad)^2 - 4ac > 0$. By using the new variable

$$u = mx \quad \text{where } m = \frac{r - ad + \sqrt{(r+ad)^2 - 4ac}}{2a}$$

Eq. (4) can be rewritten as

$$u' = \frac{a\,m\,u\,(u-\theta)\,(1-u)}{u+d/m} \quad \text{where} \quad \theta = \frac{\left(r - ad - \sqrt{(r+ad)^2 - 4ac}\right)^2}{4a(c-rd)}. \tag{5}$$

The population model (5) creates, under suitable parameter values, i.e., $(r+ad)^2 - 4ac > 0$, a threshold θ below which the population goes extinct eventually and above which the population density approaches a positive equilibrium. The simplest and generic model that captures the population dynamics of a single species with Allee effects, e.g., population model (5), can be described by

$$x' = rx\,(x-\theta)(1-x) \tag{6}$$

where r is the per capita intrinsic growth rate after rescaling and θ is a threshold that lies between 0 and 1 after rescaling. The latter, called Allee threshold, determines whether the population goes extinct or establishes itself. More precisely, the population dynamics of (6) can be summarized as follows.

Lemma 1 (Single species dynamics with Allee effects) *If the population of a single species is described by* (6)*, then it goes extinct when* $x(0) < \theta$ *while its density goes to 1 when* $x(0) > \theta$.

Thinking of model (6) as describing the population dynamics in one patch, the dynamics of two interacting identical patches with dispersal μ can be modeled by

$$x' = rx\,(x-\theta)(1-x) + \mu\,(y-x) \tag{7}$$
$$y' = ry\,(y-\theta)(1-y) + \mu\,(x-y) \tag{8}$$

where $\mu \in [0,1]$ is a dispersal parameter, representing the fraction of population migrating from one patch to another per unit of time. Although the system (7)–(8) is symmetric in x and y, asymmetry will be introduced by considering different initial conditions in each patch, i.e., $x(0) \neq y(0)$, but also in Sect. 4 by looking at the interactions between patches with different Allee thresholds. We will pay particular attention to situations where one patch is initially below and the other patch above the Allee threshold, in which case, in the absence of dispersal, the population goes extinct in the first patch but establishes itself in the second one. The main objective is to understand, based on analytical and numerical results, how the dispersal parameter μ and the Allee threshold θ affect the global dynamics, i.e., the limit sets of the system (7)–(8) and the geometry of their basin of attraction.

Our analytical results suggest the following picture of the global dynamics. Recall first that, in the absence of dispersal, the system has four locally stable equilibrium points, which correspond to cases when the population in each patch either goes extinct or gets established. In the presence of dispersal, the existence of (stable) limit cycles is also excluded: starting from almost every initial condition in $\mathbb{R}^2_+$, the system converges to an equilibrium point. This is partly proved analytically in Theorem 1 and supported by the numerical solutions of Fig. 1. Therefore, we focus our attention on the existence,

stability and basins of attraction of the equilibrium points. The effects of the dispersal parameter and the value of the Allee threshold are as follows. First, the dynamics of the deterministic two-patch model in the presence of weak dispersal are similar to that of the uncoupled system, having four locally stable equilibria (Theorem 4), i.e., the extinction state (0, 0), the expansion state (1, 1) and two asymmetric interior equilibria (x_s, y_s), (y_s, x_s), which is not retained after the inclusion of demographic stochasticity, which induces two absorbing states and two metastable states. While increasing the dispersal parameter from 0, the basin of attraction of the extinction state (0, 0) and expansion state (1, 1) increase until a certain critical value at which both patches interact enough to synchronize, which drives the system to either global extinction (0, 0) or global expansion (1, 1): there are only two attractors (Theorems 2 and 4). Above this critical value, dispersal promotes extinction when the Allee threshold is below one half but promotes survival when the Allee threshold is above one half (Theorems 2 and 3). Finally, in the presence of strong dispersal, both patches synchronize fast enough so that the global dynamics reduce to that of a single-patch model: if the initial global density, i.e., the average of the densities in both patches, is below the Allee threshold then the population goes extinct whereas if it exceeds the Allee threshold then the population expands globally (Theorem 3). In other respects, for any value of the dispersal parameter, increasing the Allee threshold promotes extinction, and populations initially below the Allee threshold in both patches are doomed to extinction, whereas populations initially above the Allee threshold in both patches expand globally. These results are stated rigorously in the following two subsections. Simulation results and detailed summary are given in the last subsection.

2.1 Global dynamics and basins of attraction

This subsection and the next one are devoted to the statement of analytical results (see Theorems 1–4) that apply to the deterministic model described by (7)–(8). In order to understand the global dynamics of the deterministic two-patch model, the first step is to identify its omega limit sets. Since the model is simply a two-dimensional ODE, its omega limit sets are either equilibrium points or limit cycles according to the Poincaré-Bendixson Theorem (Guckenheimer and Holmes 1983). As stated in the next theorem, when the dispersal parameter is sufficiently large, an application of the Dulac's criterion (Guckenheimer and Holmes 1983) reveals simple dynamics by excluding the existence of limit cycles: for any initial condition, the system converges to an equilibrium point.

Theorem 1 (Simple dynamics) *Let* $r > 0$ *and* $\theta \in (0, 1)$. *If*

$$\mu \geq r\,\theta\,(c-1) + \frac{r\,(2-c)^2\,(1+\theta)^2}{4\,(3-c)} \tag{9}$$

holds for a nonnegative number $c \in [0, 3)$, *then every trajectory of* (7)–(8) *converges to an equilibrium point.*

Theorem 1 indicates for instance that every trajectory of the system (7)–(8) converges to an equilibrium point under the condition $\mu \geq r(\theta^2 - \theta + 1)/3$ if one takes $c = 0$.

In addition, Theorem 2 below implies that if limit cycles emerge for smaller values of the dispersal parameter then each of them is included in one of the two regions of the phase space in which the population lies above the Allee threshold in one patch but below the Allee threshold in the other patch, i.e.,

$$\{(x, y) \in \Omega_0 : 0 \le x \le \theta \text{ and } \theta \le y \le 1\} \quad \text{and}$$
$$\{(x, y) \in \Omega_0 : 0 \le y \le \theta \text{ and } \theta \le x \le 1\}$$

where $\Omega_0 = \mathbb{R}^2_+$. Although the condition (9) is not a necessary condition to guarantee that the solution of (7)–(8) converges to an equilibrium point, numerical solutions (see Fig. 1) further strongly suggest that, for any value of the dispersal parameter, there is no stable limit cycle, which implies that locally stable equilibria are the only possible attractors of the system, so we focus our attention on the existence, stability and basins of attraction of the equilibrium points. We also would like to point out that if the system has no Allee effect, e.g., a metapopulation model coupled by both competition and migration with uniparental reproduction, then this system admits no periodic solutions (Proposition 1 in Gyllenberg et al. 1999).

It can be easily seen that the system (7)–(8) has three symmetric equilibria for all positive values of the parameters: one boundary equilibrium given by $E_0 = (0, 0)$ and two interior equilibria given respectively by $E_\theta = (\theta, \theta)$ and $E_1 = (1, 1)$. For obvious reasons, we call E_0 the extinction state of the system and E_1 the expansion state. Theorem 2 below indicates that, for all parameter values, these two trivial equilibria are locally stable whereas the interior equilibrium point E_θ is unstable. Hence, to understand the global dynamics of the system, the next step is to study the geometry of the basins of attraction of the two trivial equilibria, i.e.,

$$B_0 = \{(x(0), y(0)) \in \Omega_0 : \lim_{t\to\infty}(x(t), y(t)) = E_0\}$$
$$B_1 = \{(x(0), y(0)) \in \Omega_0 : \lim_{t\to\infty}(x(t), y(t)) = E_1\}.$$

Letting $\Omega_{0,\theta}$ and Ω_θ denote the subsets

$$\Omega_{0,\theta} = \{(x, y) \in \Omega_0 : 0 \le x \le \theta \text{ and } 0 \le y \le \theta\}$$
$$\Omega_\theta = \{(x, y) \in \Omega_0 : x \ge \theta \text{ and } y \ge \theta\}$$

Lemma 1 indicates that, in the absence of dispersal, the basins of attraction of E_0 and E_1 for the (uncoupled) system are given by $B_0 = \Omega_{0,\theta} \setminus E_\theta$ and $B_1 = \Omega_\theta \setminus E_\theta$. The following theorem gives some valuable insight into the geometry of the basins of attraction in the presence of dispersal.

Theorem 2 (Local stability and basins of attraction)

1. *The extinction state E_0 and expansion state E_1 are always locally stable whereas the interior fixed point E_θ is always unstable.*
2. *If $2\mu > r\theta(1-\theta)$ then E_θ is a saddle while if $2\mu < r\theta(1-\theta)$ then E_θ is a source.*

3. $\Omega_{0,\theta} \setminus E_\theta \subset B_0$. *If in addition* $\theta < 1/2$ *then*

$$B_0 \subset \{(x, y) \in \Omega_0 : x + y < 2\theta\}.$$

4. $\Omega_\theta \setminus E_\theta \subset B_1$. *If in addition* $\theta > 1/2$ *then*

$$B_1 \subset \{(x, y) \in \Omega_0 : x + y > 2\theta\}.$$

Theorem 2 indicates that the inclusion of dispersal promotes both global extinction and global expansion of the system, as the basins of attraction of both equilibrium points E_0 and E_1 are larger in the presence than in the absence of dispersal. Numerical solutions further suggest that, up to a certain critical value, increasing the dispersal parameter translates into an increase of B_0 and B_1. The value of the Allee threshold θ also plays an important role in the global dynamics. When the Allee threshold lies below one half, which seems to be common in nature (for data about Allee thresholds, see Johnson et al. 2006; Berec et al. 2007; Tobin et al. 2007; Chapter 5 in Courchamp et al. 2009), the largest possible basin of attraction of E_0 is

$$\{(x, y) \in \Omega_0 : x + y < 2\theta\}.$$

Moreover, according to the numerical solutions (see Fig. 1), increasing the Allee threshold promotes extinction of the system in the sense that, the dispersal parameter being fixed, the smaller the Allee threshold, the smaller the basin of attraction of the extinction state E_0 and the larger the basin attraction of the expansion state E_1. Finally, we would like to point out that parts 1 and 2 of the theorem hold for the system (7)–(8) but not always for two-patch models with Allee effect. A counter example is provided by the metapopulation model coupled by both competition and migration with biparental reproduction studied by Gyllenberg et al. (1999). The first step to prove parts 3 and 4 of Theorem 2 will be to identify the positive invariant sets which are included in the upper right quadrant $\Omega_0 = \mathbb{R}^2_+$. Recall that a set is called positive invariant if any trajectory starting from this set stays in this set at all future times. Since we are interested in the global dynamics of the system, our objective will be to find all the possible invariant sets in Ω_0. Notice that the union and the intersection of positive invariant sets are also positive invariant. All these positive invariant sets have an important role in understanding the dynamics in regions of the phase space where the population is below the Allee threshold in one patch but above the Allee threshold in the other patch. In particular, they will give us means of decomposing the phase space by restricting our attention to the dynamics on each invariant set and then sewing together a global solution from the invariant pieces.

2.2 Dispersal effects and multiple attractors

In this subsection, we study the effects of the dispersal parameter on the dynamics of the two-patch model when the Allee threshold is fixed. Theorem 1 suggests that the number of attractors is also equal to the number of locally stable equilibria. Our

study shows that the value of the dispersal parameter determines the number of equilibria, thus the possible number of attractors. Let S_θ denote the stable manifold of the unstable interior equilibrium E_θ, i.e.,

$$S_\theta = \{(x(0), y(0)) \in \Omega_0 : \lim_{t\to\infty} (x(t), y(t)) = E_\theta\}.$$

The following theorem indicates that, when the dispersal is sufficiently large, both patches interact enough to synchronize, which drives the system to either global extinction or global expansion: there are only two stable equilibria, the extinction state E_0 and the expansion state E_1.

Theorem 3 (Large dispersal) *Assume that*

$$\mu > \frac{r(\theta^2 - \theta + 1)}{6}. \tag{10}$$

Then, the system (7)–(8) *has only two attractors:* E_0 *and* E_1. *Moreover,*

1. *If* $\theta < 1/2$ *and* (10) *holds then*

$$\{(x, y) \in \Omega_0 \setminus S_\theta : x + y \geq 2\theta\} \subset B_1.$$

2. *If* $\theta > 1/2$ *and* (10) *holds then*

$$\{(x, y) \in \Omega_0 \setminus S_\theta : x + y \leq 2\theta\} \subset B_0.$$

If the inequality (10) holds, we can consider that the system has very strong dispersal. Then Theorem 3 indicates that, when $\theta < 1/2$, both patches synchronize fast enough so that the global dynamics reduce to the ones of a single-patch model: if the initial global density, i.e., the average of the densities in both patches, is below the Allee threshold then the population goes extinct whereas if it exceeds the Allee threshold then the population expands globally. In addition, the theoretical results in Theorems 2 and 3 suggest that the smaller the Allee threshold, the smaller the basin of attraction of the extinction state and the larger the basin of attraction of the expansion state. This agrees with the simulation results of Fig. 1.

Finally, in order to explore the number of locally stable equilibria when the dispersal is small, we now look at the nullclines of the system. Define

$$f(x) = x - \frac{rx(x - \theta)(1 - x)}{\mu}.$$

Then, the nullclines of the system (7)–(8) are given by $x = f(y)$ and $y = f(x)$. The interior equilibria are determined by the positive roots of $x = f(f(x))$, which is a polynomial with degree 9. This implies that the system has at most 8 interior equilibria since 0 is always a solution.

According to the expression of the nullclines $y = f(x)$ and $x = f(y)$ (see Fig. 7 page 964), we can see that the number of interior equilibria strongly depends upon

the value of the dispersal parameter: in the presence of strong dispersal, both patches synchronize and the system has only two positive interior equilibria E_θ and E_1, which is confirmed by Theorem 3, while in the presence of weak dispersal, there is enough independence between both patches so that the system has 8 positive interior equilibria. We are interested in the locally stable equilibria since the possible number of attractors is intimately connected to the number of stable equilibria. The following theorem gives a summary of the results related to the local and the global stability of the equilibria for different values of the parameters.

Theorem 4 (Multiple attractors)

1. *If* $r > 0$, $\theta, \mu \in [0, 1]$ *then every trajectory converges in* $[0, 1]^2$ *so all the equilibria* $(x^*, y^*) \in [0, 1]^2$.
2. *If* $6\mu > r(\theta^2 - \theta + 1)$ *then there are only three equilibria:* E_0, E_θ *and* E_1, *with* E_0 *and* E_1 *locally stable and* E_θ *saddle.*
3. *If* $4\mu < r(1-\theta)^2$ *then the nullcline*

$$y = f(x) := x - \frac{rx(x-\theta)(1-x)}{\mu}$$

has exactly two positive roots that we denote by $0 < x_1 < x_2$. *Let* $M = \max_{0 \le x \le x_1} f(x)$.

(a) If $x_1 < M < x_2$ *then the system has five fixed points with only two locally stable:* E_0 *and* E_1.

(b) If $M \ge 1$ *then the system achieves its maximum number of equilibria which is equal to 9; only four of them are locally stable: two symmetric equilibria* E_0 *and* E_1 *and two asymmetric interior equilibria* (x_s, y_s) *and* (y_s, x_s).

Part 1 of Theorem 4 suggests that we can restrict our analysis of the basins of attraction to the compact space $[0, 1]^2$. Thus, when we introduce an approximation of stochastic model in the next section, we restrict the state space to be $[0, 1]^2$. Moreover, from Theorem 4, we can see that when the dispersal parameter is small enough, the system has 9 equilibria, including four locally stable equilibria. In the presence of an Allee effect, small dispersal may promote survival: patches that are below the Allee threshold are rescued by immigrants from adjacent patches above the Allee threshold. This implies that when dispersal is introduced to a system with an Allee effect, populations can exist at intermediate densities, corresponding to the equilibria (x_s, y_s) and (y_s, x_s), as a source-sink system, or expand to high density E_1. Moreover, according to perturbation theory (Levin 1974; Amarasekare 2000), both asymmetric interior equilibria appear from the equilibria $(0, 1)$ and $(1, 0)$ of the uncoupled system, i.e., in the absence of dispersal, caused by the small perturbation μ. Therefore, we have $x_s = O(\mu)$ and $y_s = 1 - O(\mu)$. Finally, note that the absence of limit cycles given by Theorem 1 when (9) holds combined with Theorem 4 implies that

Corollary 1 (Four attractors) *If the system* (7)–(8) *has four locally stable equilibria and inequality* (9) *holds for some* $c \in [0, 3)$, *then the system has exactly four attractors.*

When the system has four attractors as stated in Corollary 1, the numerical solutions in Fig. 1 suggest that the smaller the dispersal, the smaller the basin of attraction of

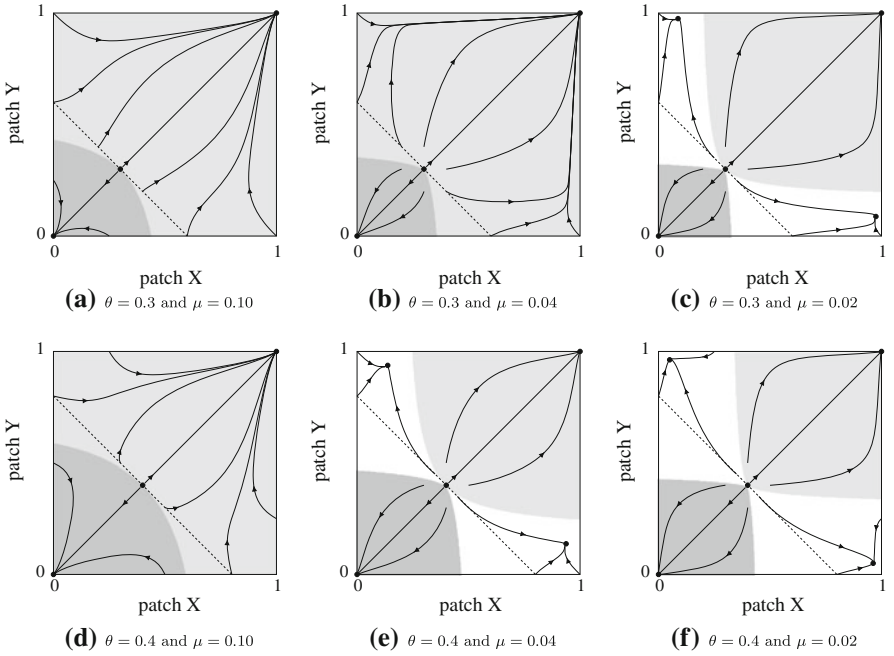


Fig. 1 Solution curves and basins of attraction of the deterministic model with $r = 1$ and $\theta < 1/2$. The values of the Allee threshold and fraction parameter, θ and μ, are indicated at the bottom of each simulation pictures. The *dark dots* are locally stable equilibria. The *solid lines* are trajectories with arrows pointing to its converging state. The *dashed line* is the straight line: $x + y = 2\theta$. The *grey* region is the basin attraction of the expansion state E_1. The *white region* is the basin attraction of (x_s, y_s) and (y_s, x_s). The *dark grey* region is the basin attraction of the extinction state E_0. Based on Theorem 4, it is enough to restrict the system (7)–(8) to the compact space $[0, 1]^2$

the extinction state and the expansion state, but the larger the basin attraction of the asymmetric interior equilibria. In particular, if $\mu \to 0$, then

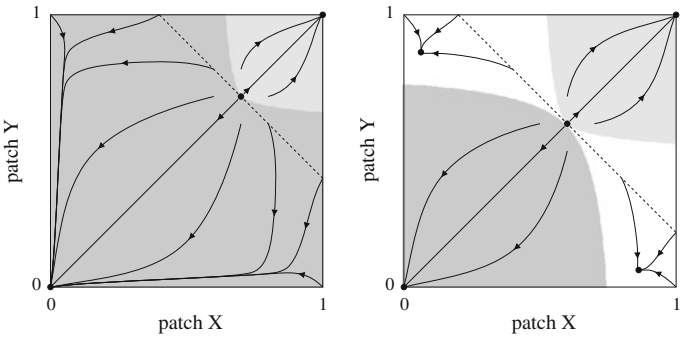
$$B_0 \to \Omega_{0,\theta}, \qquad B_1 \to \Omega_\theta, \qquad B_s \to \mathbb{R}^2_+ \setminus (\Omega_{0,\theta} \cup \Omega_\theta)$$

where B_s denotes the basin of attraction of the asymmetric interior equilibria.

2.3 Simulations and summary

Theorem 3 suggests that the larger the Allee threshold, the larger the basin of attraction of the extinction state and the smaller the basin of attraction of the expansion state when inequality (10) holds. The numerical solutions shown in Fig. 1 confirm this and give us a more complete picture of how the dispersal μ and Allee threshold θ affect the exact basin of attraction of the locally stable equilibria including asymmetric interior equilibria:

1. *Effects of dispersal* μ – Fix Allee threshold θ and growth rate r, let dispersal μ vary.



(a) $\theta = 0.7$ and $\mu = 0.04$ **(b)** $\theta = 0.8$ and $\mu = 0.04$

Fig. 2 Solution curves and basins of attraction of the deterministic model with $r = 1$ and $\theta > 1/2$. The values of the Allee threshold and fraction parameter, θ and μ, are indicated at the bottom of each simulation pictures. The *dark dots* are locally stable equilibria. The *solid lines* are trajectories with arrows pointing to its converging state. The *dashed line* is the straight line: $x + y = 2\theta$. The *grey region* is the basin attraction of the expansion state E_1. The *white* region is the basin attraction of (x_s, y_s) and (y_s, x_s). The *dark grey* region is the basin attraction of the extinction state E_0

1. When μ is small so that the system has four locally stable equilibria (μ smaller than some critical value μ_c), the smaller the dispersal, the smaller the basin of attraction of the extinction state and the expansion state, but the larger the basin of attraction of the asymmetric interior equilibria (see (e) and (f) of Fig. 1). This indicates that smaller dispersal promotes persistence of the populations in both patches by creating sink-source dynamics.
2. When μ is large so that the system has only two attractors E_0 and E_1 (μ larger than the critical value μ_c), the larger the dispersal, the larger the basin of attraction of the extinction state but the smaller the basin of attraction of the expansion state when $\theta < 1/2$ (see (a) and (b) of Fig. 1). When $\theta > 1/2$, the monotonicity is flipped due to the symmetry of the system (see Fig. 2).
3. Extreme cases: when μ is very small, the two-patch model behaves nearly like the uncoupled system, having four attractors and almost the same basins of attraction, while when $\mu/r \to \infty$, the global population behaves according to a one-patch system with Allee threshold 2θ, in particular

$$B_0 \longrightarrow \{(x, y) \in \mathbb{R}^2_+ : x + y < 2\theta\}.$$

2. *Effects of Allee threshold* θ – Fix dispersal μ and growth rate r, let Allee threshold θ vary.

   Regardless of the number of locally stable equilibria, the larger the Allee threshold, the larger the basin of attraction of the extinction state but the smaller the basin of attraction of the expansion state (see (a), (c), (d) and (f) of Fig. 1).

3. *Effects of Growth rate* r – Fix dispersal μ and Allee threshold θ, let growth rate r vary.

Table 1 Summary for the deterministic model when $\theta < 1/2$ and variations of the parameters are restricted to the case when the system has only two attractors E_0 and E_1

Parameters	Basin of attraction of E_0	Basin of attraction of E_1
Two attractors and $\theta < 1/2$		
Dispersal $\mu \uparrow$	$B_0 \uparrow$	$B_1 \downarrow$
Allee threshold $\theta \uparrow$	$B_0 \uparrow$	$B_1 \downarrow$

In particular, if $\mu/r \to \infty$ then $B_0 \to \{(x, y) \in \mathbb{R}^2_+ : x + y \leq 2\theta\}$.

Table 2 Summary for the deterministic model when $\theta < 1/2$ and variations of parameters are restricted to the case when the system has four attractors: E_0, E_1 and (x_s, y_s), (y_s, x_s)

Parameters	Basin of attr. of E_0	Basin of attr. of asymmetric equilibria	Basin of attr. of E_1
Four attractors and $\theta < 1/2$			
Dispersal $\mu \downarrow$	$B_0 \downarrow$	$B_s \uparrow$	$B_1 \downarrow$
Allee threshold $\theta \downarrow$	$B_0 \downarrow$	no monotonicity	$B_1 \uparrow$
$\mu \to 0$	$B_0 \to \Omega_{0,\theta}$	$B_s \to \mathbb{R}^2_+ \setminus (\Omega_{0,\theta} \cup \Omega_\theta)$	$B_1 \to \Omega_\theta$

By introducing the new time $\tau = t/r$, we can scale off the parameter r of the system (7)–(8) so the dispersal μ becomes μ/r. This implies that the growth rate r and the dispersal parameter μ have opposite effects on the basin of attraction of the locally stable equilibria, i.e., increasing the value of r is equivalent to decreasing the value of μ.

Tables 1, 2 give a complete picture, based on our analytical and numerical results, of how dispersal and Allee threshold affect the basin of attraction of the locally stable equilibria. We only focus on the case $\theta < 1/2$ but similar results can be deduced when $\theta > 1/2$ using the symmetry of the system (7)–(8).

3 Definition of the stochastic model and main results

While the deterministic model is similar to the one in Ackleh et al. (2007), our stochastic model differs from theirs. Their model is derived from the deterministic model by including independent Poisson increments, i.e., variability in birth, death and migration events. This gives rise to a multi-patch individual-based model for which they study numerically the probability of a successful invasion. In their article, successful invasion is defined as the event that the population size in one patch exceeds a certain threshold. However, well-known results about irreducible Markov chains imply that the population is driven almost surely to extinction which corresponds to the unique absorbing state of their stochastic process. In contrast, we model stochastically the two-patch system via a process that has two absorbing states corresponding to a global extinction and a global expansion, respectively. This allows to have a definition of successful invasion more rigorous and more tractable mathematically. In particular, while their stochastic model is designed to study numerically the probability that a popu-

lation starting near the Allee threshold in each patch gets successfully established, our model is designed to study analytically the probability that a fully occupied patch successfully invade a nearby empty patch. More precisely, to understand the effect of stochasticity on the interactions between both patches, we introduce a Markov jump process that, similarly to the deterministic model, keeps track of the evolution of the population size in each patch. To obtain a Markov process, the state is updated at random times represented by the points of a Poisson process with a certain intensity making the times between consecutive updates independent exponentially distributed random variables. Motivated by the fact that the unit square $S = [0, 1]^2$ is positive invariant for the deterministic model, we will choose this set as the state space, i.e., the state at time t is a random vector $\eta_t = (X_t, Y_t) \in S$, where the first and second coordinates represent the population size in the first and second patch, respectively. Following the deterministic model, the stochastic dynamics involve three mechanisms: expansion, extinction, and migration. To model the presence of an Allee affect, we again introduce a threshold parameter $\theta \in (0, 1)$ that can be seen as a critical size under which the population undergoes extinction and above which the population undergoes expansion, i.e., Allee threshold. This aspect is modeled by assuming that each component of the stochastic process jumps independently at rate $r > 0$ to either 0 (extinction) or 1 (expansion) depending on whether it lies below or above the Allee threshold. Recall that an event "happens at rate r" if the probability that it happens during a short time interval of length Δt approaches $r(\Delta t)$ as $\Delta t \to 0$. In particular, expansion and extinction are formally described by the conditional probabilities

$$\begin{aligned} P\left(X_{t+\Delta t} = 1 | X_t > \theta\right) &= P\left(Y_{t+\Delta t} = 1 | Y_t > \theta\right) = r\Delta t + o(\Delta t) \\ P\left(X_{t+\Delta t} = 0 | X_t < \theta\right) &= P\left(Y_{t+\Delta t} = 0 | Y_t < \theta\right) = r\Delta t + o(\Delta t). \end{aligned} \tag{11}$$

This is also equivalent to saying that the waiting time for an expansion or an extinction is exponentially distributed with mean $1/r$. Given that the population size in a given patch is at the Allee threshold, we flip a fair coin to decide whether an expansion event or an extinction event occurs at that patch which, in view of well-known properties of Poisson processes, implies that

$$\begin{aligned} P\left(X_{t+\Delta t} = 1 | X_t = \theta\right) &= P\left(Y_{t+\Delta t} = 1 | Y_t = \theta\right) = (r/2)\Delta t + o(\Delta t) \\ P\left(X_{t+\Delta t} = 0 | X_t = \theta\right) &= P\left(Y_{t+\Delta t} = 0 | Y_t = \theta\right) = (r/2)\Delta t + o(\Delta t). \end{aligned} \tag{12}$$

To understand the effects of inter-patch interactions on the evolution of the system, we also include migration events consisting of the displacement of a fraction μ of the population of each patch to the other patch. We assume that these events occur at the normalized rate 1, therefore migrations are described by

$$P\left((X_{t+\Delta t}, Y_{t+\Delta t}) = (1 - \mu)\left(X_t, Y_t\right) + \mu\left(Y_t, X_t\right)\right) = \Delta t + o(\Delta t), \tag{13}$$

We refer to Fig. 3 for a schematic illustration of the dynamics, where dark rectangles represent parts of the populations which are interchanged in the event of a migration. To analyze mathematically the stochastic process, it will be useful to look at the

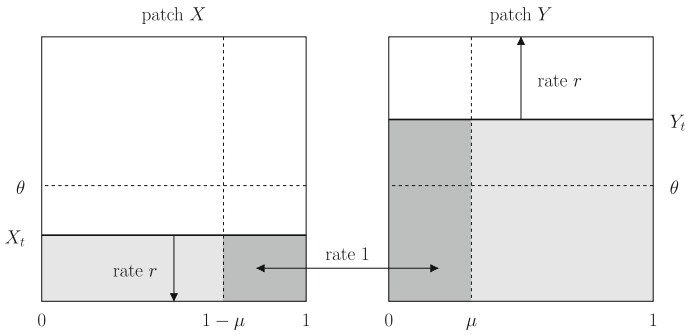


Fig. 3 Schematic representation of the stochastic model $\eta_t = (X_t, Y_t)$. The *dark rectangles* represent parts of the populations which are exchanged in the event of a migration

model as a simple example of interacting particle system. Interacting particle systems are continuous-time Markov processes whose state space maps the vertex set of a connected graph into a set representing the possible states at each vertex. The evolution is described by local interactions as the rate of change at a given vertex only depends on the configuration in its neighborhood. In particular, the Markov process $\{\eta_t\}_t$ can be seen as an interacting particle system evolving on a very simple graph that consists of only two vertices, representing both patches, connected by one edge, indicating that patches interact. The reason for looking at the stochastic model as an example of interacting particle system is that this will allow us to construct the process graphically from a collection of independent Poisson processes based on an idea of Harris (1972), which is a powerful tool to analyze the process mathematically.

We now describe in details the behavior of the process along with our main results. Note that, considering a stochastic model rather than a deterministic one, the long-term behavior is described by a set of invariant measures on the state space rather than single point equilibria. To the two trivial equilibria of the deterministic model, E_0 and E_1, correspond two invariant measures which are Dirac measures that concentrate on those two points, respectively. These two measures are two absorbing states: the configuration in which both patches are empty and the configuration in which both patches are fully occupied. We call global extinction and global expansion the events that the process eventually gets trapped into the first and the second absorbing state, respectively. Interestingly, to the two asymmetric equilibria of the deterministic model in the presence of weak dispersal correspond two quasi-stationary distributions representing two metastable states of the stochastic process (see Theorem 7): depending on the initial configuration, the transient behavior might be described by one of these two quasi-stationary distributions, but after a long random time in the presence of weak dispersal (see Theorem 6), the system gets trapped into one of the two absorbing states, suggesting that situations predicted by the deterministic model in which a small population can live next to a large population are artificially stable. Another important question is how stochasticity affects the geometry of the basins of attraction of the two absorbing states, although strictly speaking there is no basin of attraction for the stochastic model since the limiting behavior might be unpredictable, and how fast the system reaches absorption. We will see that there is a set of initial configurations for which the limiting behavior of the stochastic process is predictable, and absorption

to one of the two absorbing states occurs quickly (see Theorem 5). Starting from any other configuration, the limiting behavior becomes unpredictable in the sense that the process may reach any of the two absorbing states with positive probability. In the presence of weak dispersal, however, the limit is almost predictable in the sense that the probability that the system undergoes a global expansion after exiting one of its metastable states approaches zero or one (see Theorem 8). Whether the system gets trapped into one or the other absorbing state strongly depends on the value of the Allee threshold. The limit is less and less predictable and the time to absorption shorter and shorter as the dispersal parameter increases.

3.1 Predictable behavior

This subsection and the next one are devoted to the statement of analytical results (see Theorem 5–8) that apply to the stochastic model described by the transition rates (11)–(13). In order to describe rigorously the behavior of the stochastic model, our main objective is to estimate the times to absorption

$$\tau^+ = \inf\{t \geq 0 : X_t = Y_t = 1\} \quad \text{and} \quad \tau^- = \inf\{t \geq 0 : X_t = Y_t = 0\},$$

and the corresponding probabilities of absorption,

$$P(\tau = \tau^+) \quad \text{and} \quad P(\tau = \tau^-) \quad \text{where} \quad \tau = \min(\tau^+, \tau^-),$$

as a function of the initial configuration and the three parameters of the system. As previously explained, in contrast with the deterministic model which can have up to four distinct attractors, with probability one, either global expansion or global extinction occurs for the stochastic process, i.e.,

$$P(\tau < \infty) = P(\tau = \tau^+) + P(\tau = \tau^-) = 1.$$

The state space can be divided into four subsets. Starting from only two of these subsets the limit is predictable in the sense that

$$P(\tau = \tau^+) \in \{0, 1\}.$$

We call an upper configuration any configuration of the system in which the population size in each patch exceeds the Allee threshold, and a lower configuration any configuration in which the population size in each patch lies below the Allee threshold. These sets are denoted respectively by

$$\Omega^+ = \{(x, y) \in S : \min(x, y) > \theta\} = \Omega_{\theta,1} \setminus \{(x, y) : x = \theta \text{ or } y = \theta\}$$
$$\Omega^- = \{(x, y) \in S : \max(x, y) < \theta\} = \Omega_{0,\theta} \setminus \{(x, y) : x = \theta \text{ or } y = \theta\}.$$

Note that the set of upper configurations is closed under the dynamics, i.e., once the system hits an upper configuration, the configuration at any later time is also an upper

configuration. This implies that, starting from an upper configuration, global expansion occurs with probability one. Similarly, starting from a lower configuration, global extinction occurs with probability one. By representing the process graphically, the time to absorption can be computed explicitly, as stated in the following theorem.

Theorem 5 (time to absorption) *We have*

$$\mathbb{E}[\,\tau^+\,|(X_0, Y_0) \in \Omega^+] = \mathbb{E}[\,\tau^-\,|(X_0, Y_0) \in \Omega^-] = \frac{6r+1}{2r^2}.$$

The previous theorem indicates that, starting from an upper configuration, the system converges with probability one to the absorbing state $(1, 1)$, whereas starting from a lower configuration, it converges with probability one to the other absorbing state $(0, 0)$. This result can be seen as the analog of Theorem 2 which states that the sets of upper and lower configurations are included in the basin of attraction of the equilibrium points E_1 and E_0, respectively. Theorem 5 also indicates that, when the rates at which expansions, extinctions, and migrations occur are of the same order, the expected time to absorption is quite short.

3.2 Metastability

The long-term behavior of the process starting from a configuration which is neither an upper configuration nor a lower configuration is more difficult to study as the probabilities of global expansion and global extinction are both strictly positive, which we shall refer to as unpredictable behavior. We will prove that, in any case, the system hits either an upper or a lower configuration at a random time which is almost surely finite, after which it evolves as indicated by Theorem 5. Hence, the time to absorption and probabilities of global expansion and extinction can be determined by estimating the hitting times

$$T^+ = \inf\{t \geq 0 : (X_t, Y_t) \in \Omega^+\} \quad \text{and} \quad T^- = \inf\{t \geq 0 : (X_t, Y_t) \in \Omega^-\}$$

and the corresponding hitting probabilities

$$P(T = T^+) \quad \text{and} \quad P(T = T^-) \quad \text{where} \quad T = \min(T^+, T^-),$$

since Theorem 5 implies that

$$\mathbb{E}[\tau] = \mathbb{E}[T] + \frac{6r+1}{2r^2} \quad \text{and} \quad P(\tau = \tau^+) = P(T = T^+).$$

Even though our next results hold for all values of the parameters, they indicate that interesting behaviors emerge when the dispersal parameter μ is small. In contrast with the deterministic model which, in this case, has four attractors, as indicated by Theorem 4, the stochastic model first exhibits a metastable behavior by oscillating for an arbitrarily long time around one of the two nontrivial equilibria of the deterministic

model, and then gets trapped into one of its two absorbing states. The limit is almost predictable as the probability of global expansion approaches either 0 or 1 depending on the value of the threshold parameter. For simplicity and since the system is symmetric, we shall assume that $X_0 = 0$ and $Y_0 = 1$ but the proofs of our results easily extend to the more general case when

$$0 < \mu \ll \min\{|X_0 - \theta|, |Y_0 - \theta|\}.$$

Recall that, starting from an upper configuration or a lower configuration, the time to absorption is rather small. In contrast, when $X_0 = 0$, $Y_0 = 1$ and μ is small, the stochastic process converges to a quasi-stationary distribution in which the population size at patch X is relatively close to 0 and the population size at patch Y relatively close to 1, and stays at its quasi-stationary distribution for a very long time, i.e., the expected value of time T is large. However, due to stochasticity, the system reaches eventually an upper or a lower configuration, and then gets trapped rapidly. The next theorem gives an explicit lower bound of the expected value of the hitting time, which is the time the system stays at its quasi-stationary distribution.

Theorem 6 (metastability) *For any initial configuration, we have*

$$P(T < \infty) = P(\tau < \infty) = 1.$$

Moreover,

$$\mathbb{E}[T \mid (X_0, Y_0) = (0, 1)] \geq \frac{n_0}{2 + 4r} \left(\frac{1 + 2r}{1 + r} \right)^{n_0}$$

where

$$n_0 = \frac{1}{2} \left\lfloor \frac{\min(\ln(1 - \theta), \ln(\theta))}{\ln(1 - \mu)} \right\rfloor.$$

Note that, when the Allee threshold is bounded away from 0 and 1, and the dispersal parameter is small, n_0 is large, and so is the expected value of the hitting time T. Starting from one patch empty and the other patch fully occupied, the time to reach either an upper or a lower configuration can be divided into independent trials, each of them having a small probability of success when μ is small. In particular, when properly rescaled, the time to absorption is stochastically larger than a geometric random variable with small parameter, which is the key to proving Theorem 6. This strongly suggests that, similarly to the geometric random variable, the variance of the time to absorption is of the order of the square of its expected value. This indicates that, over a large number of independent realizations, a large fraction of these realizations will result in time to absorption much smaller than the expected value, though still very large, whereas a smaller fraction of the realizations will result in time to absorption much larger than the expected value, which is also supported by numerical simulations of the stochastic model. Now, note that,

before the hitting time, no expansion event can occur at patch X while no extinction event can occur at patch Y. To gain insight into the metastable state of the stochastic model, it is thus natural to look at the stationary distribution of the Markov process $\bar{\eta}_t = (\bar{X}_t, \bar{Y}_t)$ with state space $S = [0, 1]^2$, and whose evolution is given by

$$P\left(\bar{X}_{t+\Delta t} = 0 | \bar{X}_t \neq 0\right) = P\left(\bar{Y}_{t+\Delta t} = 1 | \bar{Y}_t \neq 1\right) = r\Delta t + o(\Delta t)$$
$$P\left((\bar{X}_{t+\Delta t}, \bar{Y}_{t+\Delta t}) = (1-\mu)\left(\bar{X}_t, \bar{Y}_t\right) + \mu\left(\bar{Y}_t, \bar{X}_t\right)\right) = \Delta t + o(\Delta t),$$

where Δt is a small time interval. In other words, the new process $\{\bar{\eta}_t\}_t$ is obtained from the original two-patch model by assuming that only extinction events at patch X and only expansion events at patch Y can occur. Note that, strictly speaking, the stationary distribution, say ν, of the new process is not the quasi-stationary distribution of the two-patch model. However, it is a good approximation of the quasi-stationary distribution when the process is metastable, i.e., when the dispersal parameter μ is very small. In particular, the following theorem gives a good approximation of the behavior before the hitting time.

Theorem 7 (metastable state) *Under the measure ν we have*

$$\mathbb{E}_\nu\left(\bar{X}_t\right) \leq \frac{\mu}{r+\mu} \quad \text{and} \quad \mathbb{E}_\nu\left(\bar{Y}_t\right) \geq 1 - \frac{\mu}{r+\mu}.$$

This indicates that, when μ is small, the population size at patch X is close to 0 (i.e., $O(\mu)$) and the population size at patch Y close to 1 (i.e., $1 - O(\mu)$). The expected values above have to be thought of as the analog of the two asymmetric equilibria of the deterministic model: (x_s, y_s) and (y_s, x_s). After evolving a long time according to its quasi-stationary distribution, the process hits either an upper or a lower configuration, so the last question we would like to answer is whether global expansion or global extinction occurs after the system exits its metastable state. Starting from an upper or a lower configuration, the answer is given by Theorem 5. Starting from $X_0 = 0$ and $Y_0 = 1$, the symmetry of the model implies that

$$P\left(\tau = \tau^+\right) = P\left(\tau = \tau^-\right) = 1/2 \quad \text{whenever } \theta = 1/2.$$

Our last result shows that, when $\theta \neq 1/2$ and $\mu > 0$ is small, the limiting behavior of the system is almost predictable in the sense that the probability of global expansion approaches either 0 or 1.

Theorem 8 (hitting probabilities) *Assume that $\theta < 1/2$. Then*

$$\frac{P\left(T = T^- | (X_0, Y_0) = (0, 1)\right)}{P\left(T = T^+ | (X_0, Y_0) = (0, 1)\right)} \leq \left(\frac{1}{1+r}\right)^{m_0}$$

where

$$m_0 = \left\lfloor \frac{\ln(2\theta)}{\ln(1-\mu)} \right\rfloor.$$

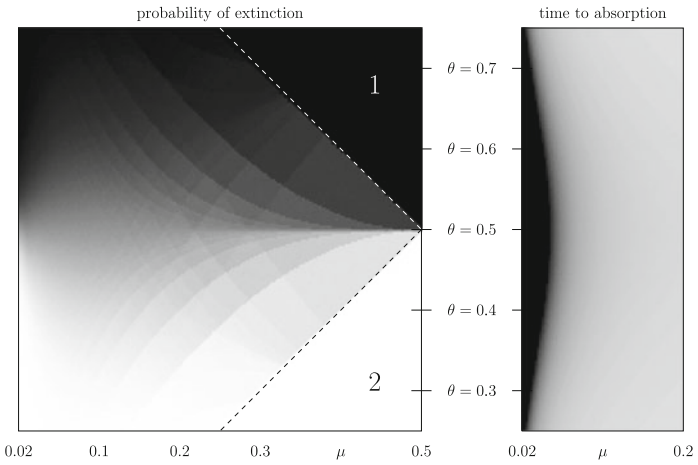
The previous theorem indicates that, when $\mu > 0$ is small,

$$\begin{aligned} P\left(\tau = \tau^+ | (X_0, Y_0) = (0,1)\right) &= P\left(T = T^+ | (X_0, Y_0) = (0,1)\right) \\ &= 1 - P\left(T = T^- | (X_0, Y_0) = (0,1)\right) \geq 1 - (1+r)^{-m_0} \approx 1. \end{aligned}$$

In particular, in contrast with the deterministic model for which the limit depends on the initial condition and the geometry of the basins of attraction, starting from any initial configuration but an upper or a lower configuration, the limiting behavior of the stochastic model is only sensitive to the value of the parameters, with the Allee threshold θ playing a central role.

3.3 Simulation results

While Theorem 5 gives an exact estimate of the time to absorption starting from particular initial conditions, the other results provide theoretical lower and upper bounds that allows us to gain a valuable insight into the long-term behavior of the stochastic two-patch model in the presence of weak dispersal. To better understand the combined effect of the Allee threshold and dispersal parameter when starting from heterogeneous initial conditions, we refer the reader to the numerical simulations of Fig. 4. The left panel of the figure represents the probability of a global extinction, with the probability increasing with the darkness, and the right panel the expected time to absorption, with time increasing with the darkness, as a function of the dispersal parameter and the Allee threshold. The tables provide some numerical values of the probability of extinction and expected time to absorption averaged over 10,000 independent realizations of the stochastic process for specific values of the parameters. The predictions based on Theorems 6 and 8 that the time to extinction blows up and the probability of extinction approaches either zero or one in the presence of weak dispersal appears clearly looking at the left side of both panels and the left column of the tables for which $\mu = 0.02$. The left picture and first table of Fig. 4 further indicate that the probability of a global extinction depends non-monotonically upon the dispersal parameter: when the Allee threshold is below one half, the probability of extinction first increases with the dispersal parameter and then decreases after the dispersal reaches a critical value that depends on θ, which can be easily seen in the row $\theta = 0.45$ of the table. When the Allee threshold exceeds one half, the monotonicity is flipped. Simulations also indicate that, the dispersal parameter being fixed, the probability of extinction increases as the Allee threshold increases. Although we omit the details of the proof, this can be easily shown analytically invoking a standard coupling argument to compare two processes, the first one with Allee threshold θ_1 and the second one with $\theta_2 > \theta_1$, the other parameters being the same for both processes. The black triangle labeled 1 delimited by the white dashed line in the upper right corner of the left picture reveals



proba.	$\mu = 0.02$	$\mu = 0.05$	$\mu = 0.10$	$\mu = 0.20$	$\mu = 0.30$	$\mu = 0.40$	$\mu = 0.50$
$\theta = 0.75$	1.000	0.999	0.995	0.999	1.000	1.000	1.000
$\theta = 0.70$	1.000	0.993	0.978	0.991	0.972	1.000	1.000
$\theta = 0.65$	0.999	0.973	0.928	0.928	0.944	1.000	1.000
$\theta = 0.60$	0.992	0.916	0.834	0.823	0.905	0.852	1.000
$\theta = 0.55$	0.910	0.759	0.681	0.719	0.779	0.857	1.000
$\theta = 0.50$	0.439	0.506	0.496	0.502	0.502	0.488	0.000
$\theta = 0.45$	0.056	0.247	0.316	0.288	0.230	0.141	0.000
$\theta = 0.40$	0.004	0.083	0.163	0.176	0.099	0.000	0.000
$\theta = 0.35$	0.000	0.030	0.066	0.078	0.062	0.000	0.000
$\theta = 0.30$	0.000	0.007	0.022	0.011	0.000	0.000	0.000
$\theta = 0.25$	0.000	0.001	0.007	0.002	0.000	0.000	0.000

time	$\mu = 0.02$	$\mu = 0.05$	$\mu = 0.10$	$\mu = 0.20$	$\mu = 0.30$	$\mu = 0.40$	$\mu = 0.50$
$\theta = 0.75$	152.711	28.888	19.749	16.133	15.165	14.994	15.058
$\theta = 0.70$	334.855	38.060	21.694	16.526	16.291	15.183	14.960
$\theta = 0.65$	815.171	52.247	24.436	17.526	16.608	15.027	15.164
$\theta = 0.60$	2052.663	70.766	27.399	18.973	16.458	16.173	15.065
$\theta = 0.55$	6058.586	91.957	29.515	19.258	16.877	16.092	14.820
$\theta = 0.50$	10520.799	102.694	30.245	20.034	18.747	18.606	14.884
$\theta = 0.45$	5566.830	91.369	29.453	19.086	16.838	15.946	14.907
$\theta = 0.40$	2075.433	70.278	27.318	18.892	16.220	14.838	14.946
$\theta = 0.35$	829.009	51.486	23.954	17.500	16.421	14.973	14.936
$\theta = 0.30$	339.504	37.717	21.467	16.309	14.867	15.143	14.993
$\theta = 0.25$	149.811	28.528	19.424	16.066	15.021	14.885	14.934

Fig. 4 Simulation results for the probability of a global extinction and the time to absorption of the stochastic model starting with one empty patch and one fully occupied patch and with growth parameter $r = 0.25$. *Left* the gradation of *grey* represents the probability of a global extinction ranging from 0 = *white* to 1 = *black*. *Right* the gradation of grey represents the time to absorption ranging from 0 = white to 100 or more = black. In both pictures, the probability and time are computed from the average of 10,000 independent simulation runs for 200 different values of the Allee threshold ranging from 0.25 to 0.75. These are further computed for 190 different values of the dispersal parameter ranging from 0.02 to 0.50, and 76 different values of the dispersal parameter ranging from 0.02 to 0.20, respectively. The tables give the values of the probability of a global extinction and the time to absorption for particular values of the parameters μ and θ

that global extinction occurs almost surely when $\theta > 1 - \mu$. Indeed, starting from the heterogeneous condition $X_0 = 0$ and $Y_0 = 1$, after the first migration event, we have

$$X_t = \mu \quad \text{and} \quad Y_t = 1 - \mu \quad \text{and so} \quad \max(X_t, Y_t) = 1 - \mu < \theta.$$

In particular, both patches are below the Allee threshold from which it follows that the population goes extinct eventually. Almost sure global expansion in the parameter region corresponding to the lower right white triangle labeled 2 can be proved similarly. Finally, as suggested by Theorem 6, the right picture and second table of Fig. 4 indicate that the expected value of the time to absorption increases as the dispersal parameter decreases but also as the Allee threshold gets closer to one half, which can again be proved analytically based on standard coupling arguments even though we omit the details of the proof.

4 Asymmetric two-patch models

In this section, we investigate the dynamics of more general asymmetric deterministic and stochastic models, focusing on the case where the Allee thresholds in both patches are different.

4.1 The asymmetric deterministic model

The asymmetric deterministic two-patch model is naturally derived from the symmetric one (7)–(8) by considering different Allee thresholds in both patches, in the following manner:

$$x' = rx\,(x - \theta_1)(1 - x) + \mu\,(y - x) \tag{14}$$
$$y' = ry\,(y - \theta_2)(1 - y) + \mu\,(x - y) \tag{15}$$

where the Allee thersholds $\theta_i \in (0, 1)$ for $i = 1, 2$ and where the parameters $r, \mu > 0$ are defined as in the system (7)–(8). To fix the ideas, we assume that $\theta_1 < \theta_2$ throughout this section, but in view of the obvious symmetry of the model, results in this case are easy to generalize to the case when $\theta_1 > \theta_2$. The main dynamics of (14)–(15) are similar to the ones of the symmetric system (7)–(8). Although we omit the proofs here, the proofs related to the symmetric model in Sect. 6 can be extended to give the following results:

1. *Positively invariant and bounded*: Starting from any initial condition in $\mathbb{R}^2_+$, the solution is attracted to the compact space $[0, 1]^2$ so the system (14)–(15) can also be restricted to the set $[0, 1]^2$.
2. *Local stability*: The system (14)–(15) always has three equilibria given by

$$E_0 = (0, 0), \qquad E_{\theta_{1,2}} = (\theta_1, \theta_2), \qquad E_1 = (1, 1),$$

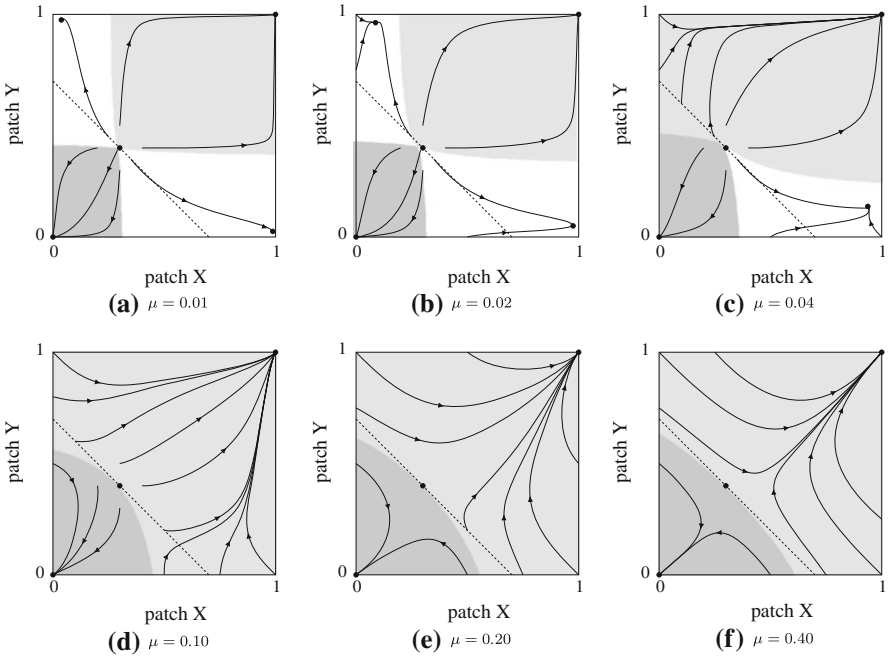


Fig. 5 Solution curves and basins of attraction when $r = 1$ and $\theta_1 = 0.3 < 0.4 = \theta_2$. The value of the dispersal parameter μ appears at the bottom of each picture. The *dashed line* is the straight line with equation $x + y = \theta_1 + \theta_2$. The *color code*, *dots* and *arrows* have the same interpretation as in Fig. 1

with E_0 and E_1 always locally stable whereas the interior fixed point $E_{\theta_{1,2}}$ is always unstable.

3. *Multiple attractors*: The number of attractors may differ from the symmetric case. More precisely, the growth rate r and Allee threshold $\theta_1 < \theta_2$ being fixed, we observe the following behavior as the dispersal parameter μ varies. As in the symmetric case, there are only two attractors, namely E_0 and E_1, for sufficiently large values of μ (see (d)-(f) in Fig. 5 for numerical results). Also, there are four attractors, the two symmetric attractors E_0 and E_1, and two asymmetric attractors

$$E_{as1} = (x_{s1}, y_{s1}) \quad \text{with} \quad x_{s1} > y_{s1} \quad \text{and} \quad E_{as1} = (x_{s2}, y_{s2}) \quad \text{with} \quad x_{s2} < y_{s2}$$

for sufficiently small values of the dispersal parameter μ (see panels (a)–(b) in Fig. 5). However, in contrast with the symmetric case, the asymmetric system has three attractors for intermediate values of the dispersal parameter μ. In the case when $\theta_1 < \theta_2$ these attractors are E_0 and E_1, and the asymmetric attractor E_{as1} (see panel (c) in Fig. 5 for a picture).

4. *Basins of attraction of E_0 and E_1*: By modifying the proof of Theorem 3, we expect the asymmetric system (14)–(15) to have only the two attractors E_0 and E_1 under a condition similar to (10). Moreover, our numerical results suggest that

Table 3 Summary for the deterministic model when $\theta_1 + \theta_2 < 1$ and variations of the parameters are restricted to the case when the system has only two attractors: E_0 and E_1

Parameters	Basin of attraction of E_0	Basin of attraction of E_1
Two attractors and $\theta_1 + \theta_2 < 1$		
Dispersal $\mu \uparrow$	$B_0 \uparrow$	$B_1 \downarrow$

In particular, if $\mu/r \to \infty$ then $B_0 \to \{(x, y) \in \mathbb{R}^2_+ : x + y \leq \theta_1 + \theta_2\}$

Table 4 Summary for the asymmetric deterministic model when $\theta_1 + \theta_2 < 1$ and variations of parameters are restricted to the case when the system has four attractors: E_0, E_1, E_{as1} and E_{as2}

Parameters	Basin of attr. of E_0	Basin of attr. of E_{s1}	Basin of attr. of E_{s2}	Basin of attr. of E_1
Four attractors and $\theta_1 + \theta_2 < 1$				
Dispersal $\mu \downarrow$	$B_0 \downarrow$	$B_{s1} \uparrow$	$B_{s2} \uparrow$	$B_1 \downarrow$
$\mu \to 0$	$B_0 \to \Omega_0$	$B_{s1} \to \Omega_{s1}$	$B_{s2} \to \Omega_{s2}$	$B_1 \to \Omega_1$

1. If $\theta_1 + \theta_2 < 1$ and the system (14)–(15) only has the two attractors E_0 and E_1, then

$$B_0 \subset \{(x, y) : x + y \leq \theta_1 + \theta_2\} \setminus S_{\theta_{1,2}}$$

where $S_{\theta_{1,2}}$ denotes the stable manifold of the unstable interior equilibrium $E_{\theta_{1,2}} = (\theta_1, \theta_2)$. We again refer to the panels (d)–(f) of Fig. 5 for an illustration.

2. If $\theta_1 + \theta_2 > 1$ and the system (14)–(15) only has the two attractors E_0 and E_1, then

$$B_1 \subset \{(x, y) : x + y \geq \theta_1 + \theta_2\} \setminus S_{\theta_{1,2}}.$$

Numerical results also suggest that for all values of the parameters, there is no stable limit cycle, which implies that locally stable equilibria are the only possible attractors of the system (14)–(15). Tables 3, 4 give a complete picture, based on numerical results, of how the dispersal parameter affects the basins of attraction of the locally stable equilibria when the asymmetric system (14)–(15) has two or four attractors. In these tables, B_{s1} and B_{s2} denote the basins of attraction of E_{as1} and E_{as2}, respectively, and

$$\begin{aligned}
\Omega_0 &:= \{(x, y) \in \mathbb{R}^2_+ : 0 \leq x \leq \theta_1 \text{ and } 0 \leq y \leq \theta_2\} \\
\Omega_{s1} &:= \{(x, y) \in \mathbb{R}^2_+ : \theta_1 \leq x \leq 1 \text{ and } 0 \leq y \leq \theta_2\} \\
\Omega_{s2} &:= \{(x, y) \in \mathbb{R}^2_+ : 0 \leq x \leq \theta_1 \text{ and } \theta_2 \leq y \leq 1\} \\
\Omega_1 &:= \{(x, y) \in \mathbb{R}^2_+ : \theta_1 \leq x \leq 1 \text{ and } \theta_2 \leq y \leq 1\}.
\end{aligned}$$

In conclusion, the main differences between the symmetric and the asymmetric systems are the number of attractors and the geometry of their basin of attraction. Whereas there can only be two or four attractors when $\theta_1 = \theta_2$, for fixed values of $r > 0$ and $\theta_1 < \theta_2$, there are two critical values $\mu_{c1} < \mu_{c2}$ that depend on the other parameters such that the (asymmetric) system has

Two attractors when $\mu > \mu_{c2}$
Three attractors when $\mu_{c1} < \mu < \mu_{c2}$
Four attractors when $\mu < \mu_{c1}$.

Numerical results suggest in addition that, in the case when $\theta_1 < \theta_2$, the basin of attraction of E_{as2} is always smaller than the basin of attraction of E_{as1} in the sense that if $(x, y) \in B_{s2}$ then $(y, x) \in B_{s1}$. They also give us some information on how the dispersal parameter affects the basins of attraction of the locally stable equilibria at each phase transition, i.e., from two to three, and from three to four attractors. Although we only focus on the case $\theta_1 + \theta_2 < 1$, similar results can be deduced from the numerical results when $\theta_1 + \theta_2 > 1$. In conclusion, our investigation indicates that the heterogenous environment, i.e., patches with different Allee thresholds, promotes more complicated dynamics, e.g., multiple attractors. For $\mu = 0.04$ and $r = 1$, by comparing panels (b) and (c) in Fig. 1 (symmetric model) for which $\theta = 0.3$ and 0.4, respectively, to panel (c) in Fig. 5 (asymmetric model) for which $\theta_1 = 0.3$ and $\theta_2 = 0.4$, we see that the dynamics of the asymmetric system are a mixture of the dynamics of the symmetric systems with $\theta = \theta_1$ and $\theta = \theta_2$.

4.2 The asymmetric stochastic model

The asymmetric stochastic two-patch model is derived from the symmetric one (11)–(13) by assigning different Allee thresholds to both patches. Migration events occur at the same rate and are described by (13). However, we now assume that expansion and extinction events occur at rate r in patch X depending on whether the population at that patch lies above or below some threshold θ_1 whereas the threshold in patch Y is given by an additional parameter θ_2. That is, expansion and extinction are formally described by the rates

$$\begin{aligned} P\,(X_{t+\Delta t} = 1 | X_t > \theta_1) &= P\,(Y_{t+\Delta t} = 1 | Y_t > \theta_2) = r\Delta t + o(\Delta t) \\ P\,(X_{t+\Delta t} = 0 | X_t < \theta_1) &= P\,(Y_{t+\Delta t} = 0 | Y_t < \theta_2) = r\Delta t + o(\Delta t). \end{aligned} \tag{16}$$

Similarly to the symmetric model, when the population size in a given patch is at the Allee threshold for that patch, we flip a fair coin to decide whether an expansion event or an extinction event occurs, which is formally described by the transition rates

$$\begin{aligned} P\,(X_{t+\Delta t} = 1 | X_t = \theta_1) &= P\,(Y_{t+\Delta t} = 1 | Y_t = \theta_2) = (r/2)\Delta t + o(\Delta t) \\ P\,(X_{t+\Delta t} = 0 | X_t = \theta_1) &= P\,(Y_{t+\Delta t} = 0 | Y_t = \theta_2) = (r/2)\Delta t + o(\Delta t). \end{aligned} \tag{17}$$

Note that the set of stationary distributions reduces, as for the symmetric model, to the two absorbing states corresponding to a global expansion and a global extinction,

respectively. Note also that when $\theta_1 \neq \theta_2$ the set of configurations in which the population lies above θ_1 in patch X and above θ_2 in patch Y, which corresponds to the set of upper configurations in the symmetric case, is no longer closed under the dynamics. In order to obtain sets which are closed under the dynamics, one has to define the sets of upper and lower configurations in the asymmetric case by setting

$$\Omega^+ = \{(x, y) \in S : \min(x, y) > \max(\theta_1, \theta_2)\}$$
$$\Omega^- = \{(x, y) \in S : \max(x, y) < \min(\theta_1, \theta_2)\}.$$

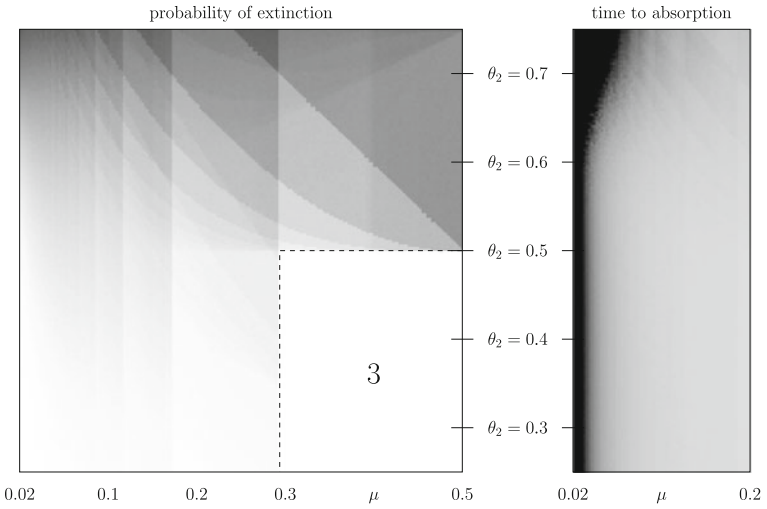
Defining in addition τ^+ and τ^- as in the symmetric case, Theorem 5 easily extends to the asymmetric model: starting from an upper configuration, global expansion occurs almost surely and the time to absorption is rather short. To understand the long-term behavior starting from a heterogeneous initial condition, we assume as before that initially patch X is empty and patch Y fully occupied, but also that $\mu \leq 1/2$. For the asymmetric model, we distinguish two cases depending on whether $\theta_1 > \theta_2$ or $\theta_1 < \theta_2$.

In the case $\theta_1 > \theta_2$, it can be proved that, starting from the initial condition mentioned above, the population size is always smaller in patch X than in patch Y before the system reaches either an upper or a lower configuration. This, together with standard coupling arguments, implies that the hitting time is stochastically larger than for the symmetric model with Allee threshold $\theta \in (\theta_2, \theta_1)$. In particular, the metastable behavior observed in the symmetric case, i.e., convergence of the time to absorption to infinity as the parameter μ tends to zero, still holds in this case. More precisely, Theorem 6 is satisfied for

$$n_0 = \frac{1}{2}\left\lfloor \frac{\min(\ln(1-\theta_1), \ln(\theta_2))}{\ln(1-\mu)} \right\rfloor > \frac{1}{2}\left\lfloor \frac{\min(\ln(1-\theta), \ln(\theta))}{\ln(1-\mu)} \right\rfloor$$
$$\text{whenever} \quad \theta \in (\theta_2, \theta_1).$$

Theorem 8 also extends to the asymmetric case: as the migration parameter μ tends to zero, the probability of a global expansion tends to one when $\theta_1 + \theta_2 < 1$ and to zero when $\theta_1 + \theta_2 > 1$. Note that when $\theta_1 = \theta_2$ one recovers the critical condition $\theta_1 = 1/2$ obtained for the symmetric model.

The asymmetric model with $\theta_1 < \theta_2$ exhibits a more interesting behavior as realizations can be divided into two very different types. Simulation results for the probability of a global extinction and the time to absorption are given in the pictures and tables of Fig. 6. The white rectangle labeled 3 corresponds to the parameters' region where $\theta_1 < \mu < 1 - \theta_2$. For such parameters, the first migration event leads the system to an upper configuration, therefore global expansion occurs almost surely. Not surprisingly, the simulation results further indicate that the same metastable behavior appears for small values of the migration parameter, with a phase transition for the probability of a global expansion around the critical point $\theta_1 + \theta_2 = 1$. However, while increasing the Allee threshold θ_2 should a priori promote global extinction in a shorter time, we observe that the time to absorption increases significantly with θ_2 for values of μ ranging from 0.03 to 0.10. This somewhat unexpected behavior follows from the fact that when $\theta_1 < \theta_2$ there is a positive probability that migration events lead the system



proba.	$\mu = 0.02$	$\mu = 0.05$	$\mu = 0.10$	$\mu = 0.20$	$\mu = 0.30$	$\mu = 0.40$	$\mu = 0.50$
$\theta_2 = 0.75$	0.714	0.688	0.665	0.580	0.571	0.590	0.495
$\theta_2 = 0.70$	0.504	0.495	0.497	0.508	0.355	0.505	0.502
$\theta_2 = 0.65$	0.285	0.359	0.366	0.323	0.306	0.481	0.498
$\theta_2 = 0.60$	0.133	0.238	0.250	0.205	0.296	0.313	0.507
$\theta_2 = 0.55$	0.047	0.166	0.170	0.152	0.178	0.311	0.495
$\theta_2 = 0.50$	0.013	0.100	0.122	0.054	0.001	0.000	0.000
$\theta_2 = 0.45$	0.003	0.057	0.088	0.036	0.000	0.000	0.000
$\theta_2 = 0.40$	0.001	0.029	0.057	0.021	0.000	0.000	0.000
$\theta_2 = 0.35$	0.000	0.015	0.039	0.018	0.000	0.000	0.000
$\theta_2 = 0.30$	0.000	0.007	0.022	0.012	0.000	0.000	0.000
$\theta_2 = 0.25$	0.000	0.002	0.012	0.007	0.000	0.000	0.000

time	$\mu = 0.02$	$\mu = 0.05$	$\mu = 0.10$	$\mu = 0.20$	$\mu = 0.30$	$\mu = 0.40$	$\mu = 0.50$
$\theta_2 = 0.75$	31764.748	232.905	49.496	29.645	22.668	21.345	18.920
$\theta_2 = 0.70$	10081.114	115.574	39.951	25.229	19.809	19.522	18.662
$\theta_2 = 0.65$	1791.596	54.449	28.729	21.308	18.483	19.330	18.657
$\theta_2 = 0.60$	322.538	39.213	24.168	19.083	18.445	17.800	18.987
$\theta_2 = 0.55$	322.666	35.855	21.821	18.059	17.185	17.511	18.861
$\theta_2 = 0.50$	332.980	35.477	20.843	16.706	15.038	14.887	15.037
$\theta_2 = 0.45$	342.961	36.709	21.047	16.209	14.998	14.906	14.831
$\theta_2 = 0.40$	338.798	37.165	21.045	16.118	15.030	15.060	14.781
$\theta_2 = 0.35$	339.458	37.934	21.570	16.280	14.914	15.044	14.948
$\theta_2 = 0.30$	338.854	38.158	21.459	16.205	15.157	15.090	14.960
$\theta_2 = 0.25$	335.733	37.936	21.310	16.134	14.875	14.805	15.107

Fig. 6 Simulation results for the probability of a global extinction and the time to absorption of the asymmetric stochastic model starting with one empty patch with Allee threshold $\theta_1 = 0.30$ and one fully occupied patch with Allee threshold θ_2 ranging from 0.25 to 0.75. The growth parameter is $r = 0.25$ for both patches. The *color code* and the meaning of the two tables are the same as in Fig. 4

to a configuration in which the population size is above θ_1 in patch X and below θ_2 in patch Y and then to the configuration in which patch X is fully occupied and patch Y empty. For such realizations, the effects of θ_1 and θ_2 are exchanged: the time to absorption is stochastically decreasing with respect to θ_1 and stochastically increasing with respect to θ_2. In particular, there are two types of realizations: realizations for which the population size is always smaller in patch X than in patch Y before hitting an upper or a lower configuration with a time to absorption shorter than in the case when $\theta_1 > \theta_2$, and realizations for which patch X becomes fully occupied and patch Y empty, which strongly contributes to increase the expected value of the time to absorption.

In conclusion, analytical and numerical results of the stochastic two-patch model can be summarized as follows. In any case, there are only two possible limiting behaviors: global expansion or global extinction. Locally stable asymmetric equilibria of the deterministic model when μ is small become metastable states for the stochastic model: as the migration parameter μ tends to zero, the time to absorption tends to infinity and the probability of a global expansion tends to one when $\theta_1 + \theta_2 < 1$ and to zero when $\theta_1 + \theta_2 > 1$, which is consistent with the symmetric case for which the phase transition occurs at $2\theta = 1$, and the metastable states are partly described by Theorem 7. In addition, we have the following dichotomy.

1. The Allee threshold θ_1 being fixed with $\theta_1 \geq \theta_2$, metastability is more pronounced (the time to absorption is larger) when the heterogeneity $\theta_1 - \theta_2$ is larger, and less pronounced in the symmetric case.
2. In contrast, the Allee threshold θ_1 being fixed with $\theta_1 < \theta_2$, metastability tends to be more pronounced when $\theta_2 - \theta_1$ is larger due to the existence of realizations for which patch X becomes fully occupied and patch Y empty. The larger $\theta_2 - \theta_1$, the larger the probability of such realizations.

Putting 1 and 2 together, we obtain that, the Allee threshold being fixed in one patch, the expected time to absorption does not depend monotonically on the Allee threshold in the other patch. Instead, the expected time seems to be increasing with respect to the absolute value $|\theta_1 - \theta_2|$, which measures the heterogeneity between both patches, although the reasons for that are quite different depending on whether $\theta_1 - \theta_2$ is positive or negative (Table 5).

5 Comparison and biological implications

Recall that, in the absence of dispersal, the deterministic model and the stochastic model result in the same predictions: the population expands in patches where the initial population size is above the Allee threshold but goes extinct in patches where the initial population size is below the Allee threshold. This induces the existence of four locally stable equilibria for the deterministic model, and four absorbing states for the stochastic model, which correspond to cases when the population in each patch either goes extinct or gets established. Including interactions between patches, our results for the symmetric deterministic model indicate that, in the presence of weak dispersal, the dynamics retain four attractors, just as in the absence of interactions, up to a

Table 5 Comparison between deterministic and stochastic models

	Deterministic model	Stochastic model
Dispersal parameter	$\theta < 1/2 \quad \vert \quad \theta > 1/2$	$\theta < 1/2 \quad \vert \quad \theta > 1/2$
No dispersal $\mu = 0$	4 Attractors	4 Absorbing states
Weak dispersal $\mu > 0$	4 Attractors	2 Absorbing states
	B_0 and $B_1 \uparrow$ as $\mu \uparrow \mu_c$	+ 2 metastable states
		P (expansion) $\approx 1 \mid P$ (extinction) ≈ 1
Critical dispersal $\mu = \mu_c$	Both patches synchronize	
	4 Attractors $\rightarrow$ 2 Attractors	2 Absorbing states
Stronger dispersal $\mu > \mu_c$	2 Attractors	2 Absorbing states
	$B_0 \uparrow$ as $\mu \uparrow \quad \vert \quad B_1 \uparrow$ as $\mu \uparrow$	Unpredictability + quick absorption
Very strong dispersal μ/r large	Same behavior as one-patch model	Same behavior as one-patch model when starting from (0, 1)

critical value μ_c when the patches synchronize: the two asymmetric equilibrium points are lost so that only global expansion and global extinction can happen. The dynamics are more complicated for the asymmetric deterministic model which has exactly three attractors for intermediate values of the dispersal parameter $\mu \in (\mu_{c1}, \mu_{c2})$. In contrast, including both stochasticity and even weak interactions, only the two absorbing states corresponding to global expansion and global extinction are retained. The most interesting behaviors emerge when the dispersal is weak, in which case, to the two asymmetric locally stable equilibria of the deterministic model, correspond two metastable states for the stochastic model.

Looking at the global dynamics, the predictions based on the analysis of the deterministic two-patch model indicate that below the critical value μ_{c1} dispersal promotes global expansion and global extinction in the sense that the basins of attraction of the two trivial fixed points expand while increasing the dispersal parameter. Above the critical value μ_{c2} dispersal promotes a global expansion when the sum of the Allee thresholds $\theta_1 + \theta_2$ exceeds one but promotes global extinction in the more realistic case when $\theta_1 + \theta_2$ lies below one. As mentioned above, in the presence of weak dispersal, both asymmetric equilibrium points become two metastable states, i.e., quasi-stationary distributions, after the inclusion of stochasticity, suggesting that situations in which a small population lives next to a large population are artificially stable: in such a context, the two-patch system evolves first as dictated by one of the two quasi-stationary distributions then, after a long random time, experiences either a global expansion or a global extinction. The Allee threshold θ_1 being fixed, analytical and numerical results suggest that the metastable behavior of the stochastic model is more pronounced as the heterogenity $|\theta_1 - \theta_2|$ increases. In addition, the long-term behavior of the stochastic model becomes almost predictable in the sense that, with very high probability, the system will undergo a global expansion when $\theta_1 + \theta_2$ lies below one and a global extinction when $\theta_1 + \theta_2$ exceeds one, which is of primary importance to predict the destiny of heterogeneous two-patch systems in the presence of weak dispersal. While increasing the dispersal parameter, the stochastic model

no longer exhibits a metastable behavior, the time to absorption decreases, and the long-term behavior becomes more and more unpredictable. In the presence of a very strong dispersal, however, the analysis of the deterministic model and the stochastic model starting from a heterogeneous configuration give the same predictions. In this case, both patches synchronize enough so that the global dynamics reduce to that of a single-patch model: if the initial global density, i.e., the average of the densities in both patches, is below $(\theta_1 + \theta_2)/2$ then the population goes extinct whereas if it exceeds $(\theta_1 + \theta_2)/2$ then the population expands globally.

Our analysis of idealized two-patch models is an important first step to understand more realistic multi-patch systems. Empirical data indicate that Allee thresholds in nature vary across species and habitat types but are typically much smaller than one half (see Johnson et al. 2006; Berec et al. 2007; Tobin et al. 2007; Chapter 5 in Courchamp et al. 2009). The predictions, based on the deterministic model in the presence of enough dispersal so that patches synchronize and on the stochastic model in the general case, that populations usually expand successfully when the Allee thresholds are small is due to the fact that only two patches interact. Literally, the critical threshold $\theta = 1/2$ in the symmetric case has to be thought of as one divided by the number of patches. Looking at a symmetric multi-patch model in which n patches interact all together, our analytical results suggest that a critical behavior should emerge for Allee thresholds near $\theta = 1/n$ when starting with a population established in only one patch, and more generally the number of patches where the population is initially established divided by the number of interacting patches. The analogous critical condition in the asymmetric case is certainly more complicated. Therefore, even for realistic values of the Allee threshold, the long-term behavior is no longer straightforward in the presence of a large number of patches. Numerical simulations can also provide a valuable insight into the long-term behavior of multi-patch models including additional refinements such as density-dependent dispersals, heterogeneous environments with possibly different growth parameters in different patches, and more importantly the inclusion of a spatial structure through a network of interactions represented by a two-dimensional regular lattice or more general planar graphs rather than a complete graph where patches interact all together.

6 Proofs for the deterministic model

Preliminary results

As explained in Sect. 2, the key to proving our main results for the determinisic model is to identify some sets which are positive invariant for the system (7)–(8). This will give us means of decomposing the phase space by restricting our attention to the dynamics on each invariant set and then sewing together a global solution from the invariant pieces. Our first preliminary result indicates that, starting from any biologically meaningful initial condition, that is any condition belonging to $\Omega_0 := \mathbb{R}^2_+$, the trajectory of the system stays in the upper right quadrant and is bounded.

Lemma 2 *The system* (7)–(8) *is positive invariant and bounded in* Ω_0.

Proof Let Γ denote the boundary of Ω_0, i.e.,

$$\Gamma = \{(x, y) \in \mathbb{R}^2 : x = 0 \text{ and } y \geq 0\} \cup \{(x, y) \in \mathbb{R}^2 : x \geq 0 \text{ and } y = 0\},$$

and observe that, if $(x(t), y(t)) \in \Gamma$, then we have the following alternative:

1. $x(t) = y(t) = 0$ and then $x(t + s) = y(t + s) = 0$ for all $s > 0$.
2. $x(t) = 0$ and $y(t) > 0$ and then $x'(t) = \mu\, y(t) > 0$.
3. $x(t) > 0$ and $y(t) = 0$ and then $y'(t) = \mu\, x(t) > 0$.

This indicates that trajectories starting from an initial condition in Ω_0 cannot cross Γ. Since in addition each trajectory is continuous, the intermediate value theorem implies that trajectories with initial condition in the quadrant Ω_0 stay in this quadrant, i.e., Ω_0 is positive invariant. The fact that the system is bounded follows from the positive invariance of Ω_0 by observing that, starting from any initial condition in Ω_0,

$$\begin{aligned} x' + y' &= rx\,(x - \theta)(1 - x) + ry\,(y - \theta)(1 - y) \\ &= r[-(x^3 + y^3) + (1 + \theta)(x^2 + y^2) - \theta(x + y)] < 0 \end{aligned}$$

whenever $x + y$ is larger than some $M(\theta) > 0$. This completes the proof. □

It follows from the proof of the previous lemma that, excluding the initial condition in which both patches are empty, the population densities in both patches are simultaneously positive at any positive time. This implies in particular that the trivial equilibrium E_0 is the only boundary equilibrium.

Lemma 3 *If* $(x(0), y(0)) \in \mathbb{R}^2_+ \setminus \{(0, 0)\}$ *then* $x(t) > 0$ *and* $y(t) > 0$ *for all* $t > 0$.

Proof By symmetry, we may assume that $x(0) > 0$ and $y(0) \geq 0$. By Lemma 2, $x(t), y(t) \in [0, M]$ at all times $t > 0$ for some constant M that depends on the initial condition. In particular,

$$x'(t) = rx\,(x - \theta)(1 - x) + \mu\,(y - x) \geq [r(x - \theta)(1 - x) - \mu]x \geq -K\,x$$

for some constant $K < \infty$, from which it follows that $x(t) > 0$ at all times $t > 0$. If $y(0) > 0$ then the same holds for $y(t)$, while if $y(0) = 0$ then

$$y'(0) = \mu\, x(0) > 0$$

and so $y(t) > 0$ for all $t \in (0, \varepsilon)$ for some small $\varepsilon > 0$. The fact that this holds at all times follows from the same reasoning as before based on the fact that both functions are bounded. □

The next lemma, which also follows from Lemma 2, is our main tool to prove Theorems 1–4. It lists some of the invariant sets of the system.

Lemma 4 *The following sets are positive invariant for the system* (7)–(8).

$$\begin{aligned}
\Omega_\theta &:= \{(x, y) \in \Omega_0 : x \geq \theta \text{ and } y \geq \theta\} \\
\Omega_1 &:= \{(x, y) \in \Omega_0 : x \geq 1 \text{ and } y \geq 1\} \\
\Omega_{0,\theta} &:= \{(x, y) \in \Omega_0 : 0 \leq x \leq \theta \text{ and } 0 \leq y \leq \theta\} \\
\Omega_{\theta,1} &:= \{(x, y) \in \Omega_0 : \theta \leq x \leq 1 \text{ and } \theta \leq y \leq 1\} \\
\Omega_{x<y} &:= \{(x, y) \in \Omega_0 : x < y\} \\
\Omega_{x>y} &:= \{(x, y) \in \Omega_0 : x > y\} \\
\Omega_{x=y} &:= \{(x, y) \in \Omega_0 : x = y\}.
\end{aligned}$$

Moreover, the dynamics along the invariant set $\Omega_{x=y}$ are described by

1. *If $x_0 = y_0 \in (0, \theta)$ then $x(t) = y(t) \to 0$ as $t \to \infty$.*
2. *If $x_0 = y_0 \in (\theta, \infty)$ then $x(t) = y(t) \to 1$ as $t \to \infty$.*

Proof The positive invariance of the set Ω_θ is a straightforward consequence of the positive invariance of the quadrant Ω_0 for the new system (u, v), where

$$u(t) = x(t) - \theta \quad \text{and} \quad v(t) = y(t) - \theta,$$

which follows from the same arguments as in the proof of Lemma 2. The positive invariance of Ω_1 can be proved similarly looking at the new system (u, v), where

$$u(t) = x(t) - 1 \quad \text{and} \quad v(t) = y(t) - 1.$$

To deal with the set $\Omega_{0,\theta}$ we first observe that if $(x(t), y(t))$ belongs to the boundary of this set then either one of the conditions 1–3 in the proof of Lemma 2 or one of the following three conditions is satisfied.

1. $x(t) = y(t) = \theta$ and then $x(t + s) = y(t + s) = \theta$ for all $s > 0$.
2. $x(t) = \theta$ and $0 \leq y(t) < \theta$ and then $x'(t) = \mu\,(y(t) - x(t)) = \mu\,(y(t) - \theta) < 0$.
3. $0 \leq x(t) < \theta$ and $y(t) = \theta$ and then $y'(t) = \mu\,(x(t) - y(t)) = \mu\,(x(t) - \theta) < 0$.

In particular, the same arguments as in the proof of Lemma 2 imply that $\Omega_{0,\theta}$ is positive invariant. The positive invariance of $\Omega_{\theta,1}$ can be proved similarly. To establish the positive invariance of the last three sets, we introduce the new functions

$$u(t) = \frac{x(t) + y(t)}{2} \quad \text{and} \quad v(t) = \frac{x(t) - y(t)}{2}.$$

A straightforward calculation shows that

$$u' = r\,u\,(u - \theta)(1 - u) + r\,v^2\,(1 + \theta - 3u) \tag{18}$$

$$v' = r\,v\,(-3u^2 + 2(1 + \theta)u - v^2 - \theta - 2\mu/r). \tag{19}$$

From (19), we see that $v = 0$ is an invariant manifold of v, which implies that the set $\Omega_{x=y}$ is positive invariant for the original system (7)–(8). Note also that if $v_0 = 0$

then (18) reduces to

$$u' = r\,u\,(u-\theta)(1-u).$$

Therefore, by applying Lemma 1, we can conclude that

$$\lim_{t\to\infty} u(t) = 0 \quad \text{when } u(0) \in (0,\theta) \quad \text{and} \quad \lim_{t\to\infty} u(t) = 1 \quad \text{when } u(0) \in (\theta,\infty),$$

which is equivalent to the last two statements of Lemma 4 in view of the definition of u and v, and the fact that $\Omega_{x=y}$ is positive invariant. Finally, since for any initial condition $u(t)$ and $v(t)$ are both bounded uniformly in time, Eq. (19) and the same argument as in the proof of Lemma 3 imply that

$$x(t)-y(t) = 2\,v(t) \geq (x_0-y_0)\,\exp(-Kt) > 0 \quad \text{for all } t>0 \text{ and for some } K<\infty.$$

The positive invariance of $\Omega_{x>y}$ follows. By symmetry, the same holds for the set $\Omega_{x<y}$. □

With Lemmas 2–4 in hands, we are now ready to prove the main results for the deterministic two-patch model described by the system (7)–(8).

Proof of Theorem 1

By Poincaré-Bendixson Theorem (Guckenheimer and Holmes 1983), the omega limit set of the system (7)–(8) is either a fixed point or a limit cycle. If the inequality (9) holds, we can use Dulac's criterion (Guckenheimer and Holmes 1983) to exclude the existence of a limit cycle. Let $c \in [0,3)$ and define the scalar function $p_c(x,y) = (xy)^{-c}$ on $\mathbb{R}^2_+$. Then,

$$\begin{aligned}
&\frac{\partial}{\partial x}[(rx\,(x-\theta)(1-x)+\mu\,(y-x))\,p_c(x,y)]\\
&\quad+\frac{\partial}{\partial y}[(ry\,(y-\theta)(1-y)+\mu\,(x-y))\,p_c(x,y)]\\
&=(xy)^{-c}[r(c-3)(x^2+y^2)+r(2-c)(1+\theta)(x+y)\\
&\quad+2r\theta(c-1)-2\mu+c\mu(2-xy^{-1}-yx^{-1})]\\
&\leq (xy)^{-c}\left[r(c-3)\left(x+\frac{(2-c)(1+\theta)}{2(c-3)}\right)^2-\frac{r(2-c)^2(1+\theta)^2}{4(c-3)}\right.\\
&\quad\left.+r(c-3)\left(y+\frac{(2-c)(1+\theta)}{2(c-3)}\right)^2-\frac{r(2-c)^2(1+\theta)^2}{4(c-3)}+2r\theta(c-1)-2\mu\right].
\end{aligned}$$

In particular, if (9) holds then the equation above is strictly negative for any $(x,y) \in \Omega \setminus E_0$. Therefore, by Dulac's criterion, the system has no limit cycle, i.e., any trajectory of (7)–(8) starting with a nonnegative initial condition converges to a fixed point. □

Proof of Theorem 2

Parts 1 and 2 of Theorem 2 about the local stability of the three symmetric equilibria follow from the analysis of the Jacobian matrices. For each of the three equilibria, we have

1. E_0 – The Jacobian matrix associated with this equilibrium is

$$J_0 = \begin{pmatrix} -r\theta - \mu & \mu \\ \mu & -r\theta - \mu \end{pmatrix} \tag{20}$$

with eigenvalues $\lambda_1 = -r\theta$ and $\lambda_2 = -r\theta - 2\mu$ associated with $(1, 1)$ and $(-1, 1)$ as their eigenvectors, respectively. We can easily conclude that the trivial boundary equilibrium E_0 is locally stable since both eigenvalues of (20) are negative.

2. E_θ – The Jacobian matrix associated with this equilibrium is

$$J_\theta = \begin{pmatrix} r\theta(1-\theta) - \mu & \mu \\ \mu & r\theta(1-\theta) \end{pmatrix} \tag{21}$$

with eigenvalues $\lambda_1 = r\theta(1-\theta)$ and $\lambda_2 = r\theta(1-\theta) - 2\mu$ associated with $(1, 1)$ and $(-1, 1)$ as their eigenvectors, respectively. We can easily conclude that the equilibrium E_θ is always unstable on the invariant set $\Omega_{x=y}$. Moreover, if $2\mu > r\theta(1-\theta)$ then E_θ is a saddle, while if $2\mu < r\theta(1-\theta)$ then E_θ is a source.

3. E_1 – The Jacobian matrix associated with this equilibrium is

$$J_1 = \begin{pmatrix} -r(1-\theta) - \mu & \mu \\ \mu & -r(1-\theta) - \mu \end{pmatrix} \tag{22}$$

with two negative eigenvalues $\lambda_1 = -r(1-\theta) < 0$ and $\lambda_2 = -r(1-\theta) - 2\mu < 0$ since $\theta < 1$. Therefore, the equilibrium E_1 is also locally stable.

To prove the third part of the theorem, we first define the function $u(t) = x(t) + y(t)$. Then

$$u' = x' + y' = rx\,(x-\theta)(1-x) + ry\,(y-\theta)(1-y).$$

To prove that

$$\Omega_{0,\theta} \setminus \{(\theta, \theta)\} \subset B_0 \tag{23}$$

we first assume that

$$(x_0, y_0) \in \Omega_{0,\theta} \setminus \{E_0, E_\theta\}.$$

Since the set $\Omega_{0,\theta}$ is positive invariant, we have $u'(t) \leq 0$ for all $t \geq 0$. Using in addition that

$$u'(t) = 0 \quad \text{if and only if} \quad u(t) = 0,$$

we can conclude that $u(t)$ converges to zero. Recalling the definition of u and invoking again the positive invariance of $\Omega_{0,\theta}$, we can deduce that $x(t)$ and $y(t)$ converge to zero so (23) holds. To prove that

$$\Omega_\theta \setminus \{(\theta, \theta)\} \subset B_1 \tag{24}$$

we now assume that

$$(x_0, y_0) \in \Omega_\theta \setminus \{E_\theta, E_1\}.$$

Then, we have the following alternative.

1. $(x_0, y_0) \in \Omega_{\theta,1} \setminus \{E_\theta, E_1\}$. Since $\Omega_{\theta,1}$ is positive invariant, we may use the same argument as before to see that the derivative of u is nonnegative and the system converges to the equilibrium point E_1.
2. $(x_0, y_0) \in \Omega_1 \setminus \{E_1\}$. Repeating again the same argument but with the positive invariant set Ω_1 implies that the system converges to E_1.
3. $(x_0, y_0) \in \Omega_\theta \setminus (\Omega_{\theta,1} \cup \Omega_1)$. We may assume that $x_0 < y_0$ without loss of generality since the system is symmetric. Then, using the positive invariance of the set $\Omega_{x<y}$ we have $x(t) < y(t)$ for all $t \geq 0$ so

$$\begin{aligned} x' &= rx\,(x-\theta)(1-x) + \mu\,(y-x) > 0 \quad \text{if } x \leq 1 \\ y' &= ry\,(y-\theta)(1-y) + \mu\,(x-y) < 0 \quad \text{if } y \geq 1. \end{aligned}$$

   This indicates that the trajectory starting at (x_0, y_0) can only exit the infinite rectangle $[\theta, 1] \times [1, \infty)$ by crossing its bottom or right side. Therefore, we have the following three possibilities.

   a. No exit: $(x(t), y(t)) \notin \Omega_{\theta,1} \cup \Omega_1$ for all $t \geq 0$. In this case, the sign of the derivatives implies convergence to E_1.
   b. Bottom side: $(x(t), y(t)) \in \Omega_{\theta,1}$ for some time $t \geq 0$. In this case, point 1 above implies convergence to the equilibrium point E_1.
   c. Right side: $(x(t), y(t)) \in \Omega_1$ for some time $t \geq 0$. In this case, point 2 above implies convergence to the equilibrium point E_1.

Combining 1-3 above implies (24). Now assume that $\theta < 1/2$. Defining

$$u(t) = \frac{x(t) + y(t)}{2} \quad \text{and} \quad v(t) = \frac{x(t) - y(t)}{2}$$

recall that the system (7)–(8) can be rewritten as

$$\begin{aligned} u' &= r\,u\,(u-\theta)(1-u) + r\,v^2\,(1+\theta-3u) \\ v' &= r\,v\,(-3u^2 + 2(1+\theta)u - v^2 - \theta - 2\mu/r). \end{aligned}$$

To prove that

$$B_0 \subset \{(x, y) \in \Omega_0 : x + y < 2\theta\} \tag{25}$$

it suffices to prove that

$$x_0 + y_0 \geq 2\theta \quad \text{implies} \quad x(t) + y(t) \geq 2\theta \quad \text{for all } t \geq 0. \tag{26}$$

Assume by contradiction that (26) is not satisfied. Then, there exists an initial condition with $x_0 + y_0 \geq 2\theta$ and a time $T > 0$ such that

$$x(0) + y(0) = x_0 + y_0 \geq 2\theta \quad \text{and} \quad x(T) + y(T) < 2\theta.$$

By continuity of the trajectories, the intermediate value theorem implies the existence of a time $t < T$ such that $u(t) = 2\theta$ therefore

$$S := \sup\{t < T : u(t) = 2\theta\}$$

is well defined and $u(t) < 2\theta$ for all $t \in (S, T]$. To prove that this leads to a contradiction, we consider the following two cases.

1. If $x_0 \neq y_0$, the invariance of $\Omega_{x<y}$ and $\Omega_{x>y}$ implies that $x(t) \neq y(t)$ at any time, from which it follows that

$$u'(S) = r\, v^2(S)\,(1 + \theta - 3\theta) = (r/4)\,(x(S) - y(S))^2(1 - 2\theta) > 0.$$

   In particular, there exists $\varepsilon > 0$ such that $u(t) > 2\theta$ for all $t \in (S, S + \varepsilon)$, which contradicts the existence of time S.
2. If $x_0 = y_0$ the result directly follows from the fact that $\Omega_\theta \cap \Omega_{x=y}$ is positive invariant, as it is the intersection of two invariant sets.

Combining 1 and 2 above yields (25). The proof of the last inclusion in Theorem 2 follows from similar arguments. □

Proof of Theorem 3

Define as previously the functions

$$u(t) = \frac{x(t) + y(t)}{2} \quad \text{and} \quad v(t) = \frac{x(t) - y(t)}{2}.$$

Using (19) above, we obtain

$$v' = r\, v \left(-3 \left(u - \frac{1+\theta}{3} \right)^2 - \frac{2}{r} \left(\mu - \frac{r(\theta^2 - \theta + 1)}{6} \right) - v^2 \right).$$

Assume first that $x_0 > y_0$. Using the fact that the set $\Omega_{x>y}$ is positive invariant by Lemma 4, we deduce that $x(t) > y(t)$ at all positive times t. In particular, recalling the definition of v, and using the expression of the derivative v' above and the fact that (10) holds, we obtain that

$$v(0) > 0 \quad \text{implies} \quad v(t) > 0 \text{ and } v'(t) < 0 \text{ for all } t \geq 0.$$

Since $v' = 0$ if and only if $v = 0$, we deduce that $v(t)$ converges to 0. By symmetry, the same can be proved of the system starting with any initial conditions such that $x_0 < y_0$. Since $\Omega_{x=y}$ is positive invariant, we have the same conclusion when the initial condition satisfies $x_0 = y_0$, which can also be seen from the expression of the derivative v'. Therefore, if (10) holds then

$$\text{for all} (x_0, y_0) \in \mathbb{R}_+^2, \quad \lim_{t\to\infty} v(t) = 0,$$

so for any $(x_0, y_0) \in \mathbb{R}_+^2$ and any $\epsilon > 0$, there exists $k > 0$ such that

$$|x(t) - y(t)| < \epsilon \qquad \text{for all } t > k.$$

It follows that any trajectory of the system converges to one of the symmetric equilibria E_0, E_θ or E_1. Now, observing that

$$\frac{r(\theta^2 - \theta + 1)}{6} - \frac{r\theta(1-\theta)}{2} = \frac{r(2\theta - 1)^2}{6} \geq 0,$$

we obtain that

$$\mu > \frac{r(\theta^2 - \theta + 1)}{6} \geq \frac{r\theta(1-\theta)}{2}.$$

In view of the expression of the Jacobian matrix (21), this implies that the equilibrium (θ, θ) is a saddle with unstable manifold

$$\Omega_{x=y} \setminus \{E_0, E_\theta, E_1\}.$$

In particular, it follows from Hartman-Grobman Theorem and the second part of Lemma 4 that there are only two attractors: E_0 and E_1. Hence, the system starting from any initial condition not belonging to the manifold S_θ converges to either E_0 or E_1. To conclude the proof, it suffices to observe that, by (26) in the proof of Theorem 2, if $\theta < 1/2$ then the set

$$\{(x, y) \in \Omega_0 \setminus S_\theta : x + y \geq 2\theta\}$$

is positive invariant so the system starting from any initial condition in this set converges to E_1, the only attractor in this invariant set. Similarly, the last statement follows

from the fact that, if $\theta > 1/2$ then the set

$$\{(x, y) \in \Omega \setminus S_\theta : x + y \leq 2\theta\}$$

is positive invariant and contains only one attractor: E_0. □

Proof of Theorem 4

We first prove that all the equilibria of the system (7)–(8) belong to the unit square $[0, 1] \times [0, 1]$. Since the system is symmetric and, by Lemma 4, the omega limit set of any initial condition $(x_0, y_0) \in \Omega_{0,\theta} \cup \Omega_\theta$ belongs to the unit square (either E_0, E_θ or E_1), it suffices to focus on the case

$$(x_0, y_0) \in \{(x, y) \in \Omega_0 : 0 \leq x \leq \theta \leq y\}.$$

Since $x_0 = y_0$ only happens when starting from (θ, θ), to avoid trivialities, we shall assume in addition that $x_0 < y_0$. Then, applying Lemma 4, we obtain that $x(t) < y(t)$ for all $t \geq 0$ which, together with (8), implies that

$$y' = ry\,(y - \theta)(1 - y) + \mu\,(x - y) < 0 \quad \text{if} y \geq 1. \tag{27}$$

Excluding the trivial case when the initial condition belongs to the stable manifold of (θ, θ), in which case its omega limit set reduces to (θ, θ), we have the following alternative.

1. If $(x(t), y(t)) \in \Omega_\theta$ for some $t \geq 0$ then, by the second part of Theorem 2, the omega limit set of the initial condition is E_1.
2. If $(x(t), y(t)) \notin \Omega_\theta$ for all $t \geq 0$ then (27) and the fact that $\Omega_{x<y}$ is positive invariant imply that $x(T) < y(T) \leq 1$ for some $T > 0$. Using as previously the continuity of the trajectories and the fact that

$$y' = \mu\,(x - y) < 0 \quad \text{if } y = 1$$

allows to invoke the intermediate value theorem and prove by contradiction that $x(t) < y(t) \leq 1$ at any time $t \geq T$.

This establishes the first part of Theorem 4. The second part follows directly from the proof of Theorem 3. To prove the third part, we first observe that the equation of the nullcline $y = f(x)$ can be rewritten as

$$y = \frac{rx}{\mu}\left[\left(x - \frac{1+\theta}{2}\right)^2 - \frac{(1-\theta)^2}{4} + \frac{\mu}{r}\right].$$

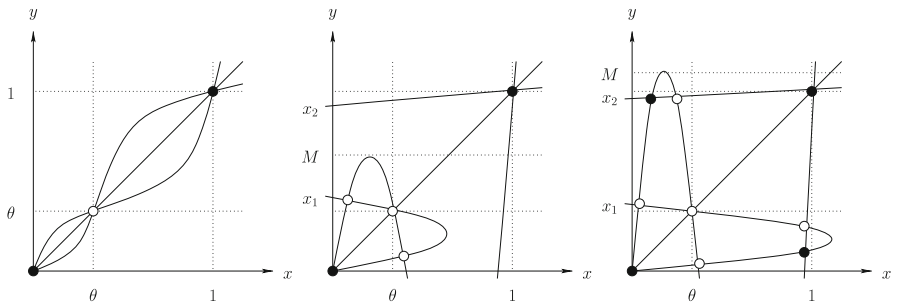


Fig. 7 Schematic presentations of nullclines of the system with black dots representing locally stable equilibria. In the rightmost two pictures, x_1 and x_2 are the only two roots of $f(x) = 0$ such that $\theta < x_1 < x_2 < 1$

In particular, if $(1-\theta)^2 > 4\mu/r$ then the nullcline intersects the x-axis at the three points with coordinates $(0, 0)$, $(x_1, 0)$ and $(x_2, 0)$ where

$$x_1 = \frac{1+\theta}{2} - \sqrt{\frac{(1-\theta)^2}{4} - \frac{\mu}{r}}$$
$$x_2 = \frac{1+\theta}{2} + \sqrt{\frac{(1-\theta)^2}{4} - \frac{\mu}{r}}.$$

Finally, a phase-plane analysis based on Fig. 7 shows that

1. If $x_1 < M < x_2$ then the system has five fixed points with only two locally stable equilibria: E_0 and E_1.
2. If $M \geq 1$ then the system achieves maximum number of equilibria which is nine, with only four locally stable equilibria.
3. If the system has less than nine equilibria, then it has only two local stable equilibria: E_0 and E_1.

Notice that $f(x)$ is a polynomial with degree 3. Moreover, provided $4\mu < r(1-\theta)^2$, it has two distinct roots $0 < x_1 < x_2$. Thus M can be computed explicitly: $M = f(x^*)$ where

$$x^* = \frac{1}{3}\left((1+\theta) - \sqrt{\theta^2 - \theta + 1 - 3\mu/r}\right)$$

is the smallest root of the quadratic function

$$f'(x) = 1 + \frac{r\theta}{\mu} - \frac{2r(1+\theta)x}{\mu} + \frac{3rx^2}{\mu}.$$

This completes the proof. □

7 Proofs for the stochastic model

Proof of Theorem 5

The first step is to prove that the set of upper configurations is closed under the dynamics. We observe that, condition on the event that the configuration is an upper configuration, only expansions and migrations can occur. Furthermore, migration events can only result in an increase of the lowest density and a decrease of the highest density, i.e., if a migration event occurs at time t and the configuration at time $t - \Delta t$ is an upper configuration then

$$\begin{aligned}&\min(X_{t-\Delta t}, Y_{t-\Delta t}) \le \min(X_t, Y_t)\\ &\quad \le \max(X_t, Y_t) \le \max(X_{t-\Delta t}, Y_{t-\Delta t}).\end{aligned}$$

It follows that the set of upper configurations (and similarly the set of lower configurations) is closed under the dynamics, i.e.,

$$\begin{aligned}&P\left((X_t, Y_t) \in \Omega^+ | (X_0, Y_0) \in \Omega^+\right)\\ &= P\left((X_t, Y_t) \in \Omega^- | (X_0, Y_0) \in \Omega^-\right) = 1\end{aligned}$$

for all times t. Since, starting from an upper configuration, the system jumps to $(1, 1)$ whenever two expansion events at X and Y occur consecutively (they are not separated by a migration event), we deduce that the stopping time τ^+ is almost surely finite. The same holds for the stopping time τ^- when starting from a lower configuration. Hence,

$$\begin{aligned}&P\left(\tau = \tau^+ < \infty | (X_0, Y_0) \in \Omega^+\right)\\ &= P\left(\tau = \tau^- < \infty | (X_0, Y_0) \in \Omega^-\right) = 1.\end{aligned}$$

To compute the expected value of the time to absorption, we now construct the stochastic process graphically from a collection of Poisson processes, relying on an idea of Harris (1972). Two Poisson processes, each with parameter r, are attached to each of the patches X and Y, and an additional Poisson process with parameter one is attached to the edge connecting the patches. All three processes are independent. Let

$$\Gamma_X = \{T_n^X : n \ge 1\}, \quad \Gamma_Y = \{T_n^Y : n \ge 1\}, \quad \Gamma_e = \{T_n^e : n \ge 1\}$$

denote these Poisson processes. At any time of the process Γ_X the population size at patch X jumps to either 0 or 1 depending on whether it is smaller or larger than θ by this time, respectively. The evolution at patch Y is defined similarly but using the Poisson process Γ_Y. At each time in Γ_e, a fraction μ of the population at each patch is displaced to the other patch. To compute the expected value, we let $t \ge 0$ and introduce the stopping times

$$T_Z = \min\{\Gamma_Z \cap (t, \infty)\} \quad \text{for } Z = X, Y, e.$$

Then, $P\,(\max(T_X, T_Y) < T_e)$ is the probability that two consecutive migration events are separated by at least one extinction-expansion event at patch X and one extinction-expansion event at patch Y. To compute this probability, we first observe that T_X and T_Y are independent exponentially distributed random variables with parameter r, from which it follows that

$$P\,(\max(T_X, T_Y) < u) = P\,(T_X < u,\ T_Y < u) = (1 - \exp(-ru))^2.$$

Since T_e is exponentially distributed with parameter 1,

$$P\,(\max(T_X, T_Y) < T_e) = \int_0^\infty \int_u^\infty e^{-v} \frac{d}{du}\left((1 - \exp(-ru))^2\right) dv\, du$$

$$= \int_0^\infty e^{-u} \frac{d}{du}\left((1-\exp(-ru))^2\right) du = \frac{2r}{r+1} - \frac{2r}{2r+1} = \frac{2r^2}{(r+1)(2r+1)} := p_s.$$

Hence, the last time a migration event occurs before absorption is equal in distribution to T^e_{J-1} where the random variable J is geometrically distributed with parameter p_s from which we deduce that

$$\mathbb{E}\,[\,\tau^+ | \theta < X_0, Y_0 < 1] = \mathbb{E}\,[T_e] \times \mathbb{E}\,[J-1] + \mathbb{E}\,[\max(T_X, T_Y)]$$

$$= \frac{(r+1)(2r+1)}{2r^2} - 1 + \int_0^\infty P\,(\max(T_X, T_Y) > u)\, du$$

$$= \frac{3r+1}{2r^2} + \int_0^\infty 1 - (1 - \exp(-ru))^2\, du = \frac{6r+1}{2r^2}.$$

The same holds for the stopping time τ^- when starting the process from a lower configuration. This completes the proof of Theorem 5. □

Proof of Theorem 6

We first prove that $P\,(T < \infty) = P\,(\tau < \infty) = 1$. Let $\epsilon > 0$ small. Then, for almost all realizations of the process, there exists an increasing sequence of random times $T_1 < \cdots < T_i < \cdots$ such that

$$\lim_{i\to\infty} T_i = \infty \quad \text{and} \quad |X_{T_i} + Y_{T_i} - 2\theta| > \epsilon \quad \text{for all } i \geq 1.$$

Moreover, there exists $K < \infty$ that does not depend on i such that, if after T_i a sequence of K migration events occur before any expansion or extinction events then the system hits either an upper configuration or a lower configuration. Since K is finite, such an event has a strictly positive probability, so the Borel-Cantelli Lemma

implies that the process hits either an upper configuration or a lower configuration after a random time which is almost surely finite: $P(T < \infty) = 1$. Theorem 5 then implies that

$$\begin{aligned} P(\tau < \infty) &= P(\tau^+ < \infty) + P(\tau^- < \infty) \\ &\geq P(\tau^+ < \infty | (X_0, Y_0) \in \Omega^+) P(T^+ < \infty) \\ &\quad + P(\tau^- < \infty | (X_0, Y_0) \in \Omega^-) P(T^- < \infty) \\ &= P(T^+ < \infty) + P(T^- < \infty) = P(T < \infty) = 1. \end{aligned}$$

To estimate the expected value of T, we observe that the transition rates of the process indicate that if at time t exactly n migration events but neither expansion nor extinction events have occurred then

$$(X_t, Y_t) = f^n(0, 1) \quad \text{where} \quad f(a, b) = (1 - \mu)(a, b) + \mu(b, a),$$

so that $X_t \leq u_n$ and $Y_t \geq v_n$ where u_n and v_n are defined recursively by

$$\begin{aligned} u_{n+1} &= (1 - \mu) u_n + \mu \quad \text{with} \quad u_0 = 0, \\ v_{n+1} &= (1 - \mu) v_n \quad \text{with} \quad v_0 = 1. \end{aligned}$$

A straightforward calculation shows that

$$u_n = \sum_{k=0}^{n-1} (u_{k+1} - u_k) = \sum_{k=0}^{n-1} (1 - \mu)^k (u_1 - u_0) = 1 - (1 - \mu)^n$$

and $v_n = 1 - u_n = (1 - \mu)^n$, therefore

$$\begin{aligned} u_n > \theta \quad &\text{if and only if } n > n_1 := \lfloor \ln(1 - \theta) / \ln(1 - \mu) \rfloor \\ v_n < \theta \quad &\text{if and only if } n > n_2 := \lfloor \ln(\theta) / \ln(1 - \mu) \rfloor \end{aligned}$$

where $\lfloor \cdot \rfloor$ is for the integer part. Now, let $\{(U_t, V_t)\}_t$ be the Markov process with state space

$$E = \{(u_i, v_j) : i, j \geq 0\}$$

and transition rates

$$\begin{aligned} P((U_{t+\Delta t}, V_{t+\Delta t}) = (0, v_j) | (U_t, V_t) = (u_i, v_j)) &= r \Delta t + o(\Delta t) \\ P((U_{t+\Delta t}, V_{t+\Delta t}) = (u_i, 1) | (U_t, V_t) = (u_i, v_j)) &= rh + o(h) \\ P((U_{t+\Delta t}, V_{t+\Delta t}) = (u_{i+1}, v_{j+1}) | (U_t, V_t) = (u_i, v_j)) &= \Delta t + o(\Delta t) \end{aligned}$$

and starting at $(U_0, V_0) = (0, 1)$. We call W-, N-, and SE-jumps, the jumps described by the three transition rates above, respectively, and refer the reader to the left-hand side of Fig. 8 for an illustration of the process.

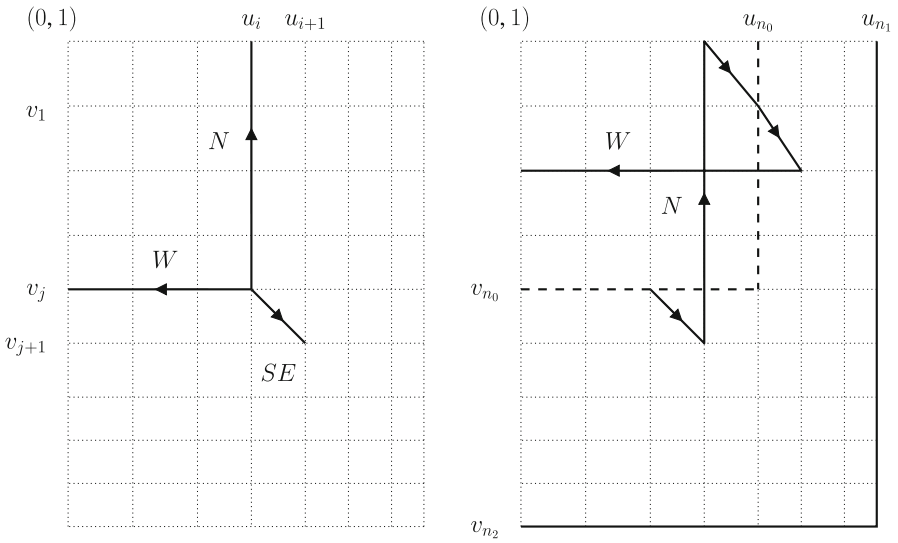


Fig. 8 Schematic representation of the stochastic process (U_t, V_t)

By construction of the sequences $(u_n)_n$ and $(v_n)_n$, we have

$$P\,(X_t \geq a | T > t) \leq P\,(U_t \geq a)$$
$$P\,(Y_t \geq a | T > t) \geq P\,(V_t \geq a)$$

for all $a \in [0, 1]$, i.e., before the process hits an upper or a lower configuration, X_t is stochastically smaller than U_t while Y_t is stochastically larger than V_t. This implies that $\mathbb{E}\,[T] \geq \mathbb{E}\,[T^*]$ where

$$T^* = \inf\{t \geq 0 : U_t > u_{n_1} \text{or } V_t < v_{n_2}\}.$$
$$E_1 = \{(u_i, v_j) : i, j \leq n_0\} \quad \text{and} \quad E_2 = \{(u_i, v_j) : i \leq n_1 \text{ and } j \leq n_2\}.$$

Then, T^* is the first time (U_t, V_t) exits the set E_2, i.e.,

$$T^* = \inf\{t \geq 0 : (U_t, V_t) \notin E_2\}.$$

So, to bound $\mathbb{E}\,[T^*]$ from below, it suffices to prove that $(U_t, V_t) \in E_2$ for an arbitrarily long time. The idea is to prove that, when starting from the smaller rectangle E_1, the process stays in E_2 and comes back to E_1 after n_0 jumps with probability close to 1. Using in addition the Markov property, we obtain that the number of jumps required to exit E_2 is stochastically larger than n_0 times a geometric random variable with small success probability. To make this argument precise, we let $(\mathcal{U}_n, \mathcal{V}_n)$ denote the embedded discrete-time Markov chain associated with the process (U_t, V_t). To count the number of steps needed to exit the rectangle E_2, we define a sequence of

Bernoulli random variables $\{Z_k : k \geq 1\}$ associated to $(\mathcal{U}_n, \mathcal{V}_n)$ by setting

$$Z_k = \begin{cases} 0 \text{ if there is at least one } N\text{-jump and one } W\text{-jump between times} \\ \quad (k-1)n_0 + 1 \text{ and } kn_0 \\ 1 \text{ if there is no } N\text{-jump or no } W\text{-jump between times } (k-1)n_0+1 \text{ and } kn_0 \end{cases}$$

Since $(\mathcal{U}_n, \mathcal{V}_n)$ is a discrete-time Markov chain, the random variables Z_k are independent Bernoulli random variables, and a straightforward calculation shows that the success probability is given by

$$P(Z_k = 1) \leq 2\left(1 - \frac{r}{1+2r}\right)^{n_0} = 2\left(\frac{1+r}{1+2r}\right)^{n_0}.$$

Moreover, since $n_0 = (1/2)\ \min(n_1, n_2)$, we have that

$$\begin{array}{l} (\mathcal{U}_{kn_0}, \mathcal{V}_{kn_0}) \in E_1 \text{ and } Z_{k+1} = 0 \quad \text{implies that} \\ (\mathcal{U}_n, \mathcal{V}_n) \in E_2 \text{ for all } kn_0 \leq n \leq (k+1)n_0 \text{ and } (\mathcal{U}_{(k+1)n_0}, \mathcal{V}_{(k+1)n_0}) \in E_1. \end{array}$$

See the right-hand side of Fig. 8. This indicates that

$$Z_1 = Z_2 = \cdots = Z_k = 0 \quad \Longrightarrow \quad (\mathcal{U}_n, \mathcal{V}_n) \in E_2 \quad \text{for all } n \leq kn_0.$$

Finally, using that (U_t, V_t) jumps at rate $1 + 2r$ and that $\inf\{k : Z_k = 1\}$ is stochastically larger than a geometric random variable Z with success probability $P(Z_k = 1)$ we can conclude that

$$\mathbb{E}[T] \geq \mathbb{E}[T^*] \geq \frac{n_0}{1+2r}\mathbb{E}[Z] = \frac{n_0}{2+4r}\left(\frac{1+2r}{1+r}\right)^{n_0}.$$

This completes the proof. □

Proof of Theorem 7

First, we observe that the process U_t introduced in the proof of Theorem 6 is stochastically larger than $\bar{X}_t$ so to prove the first inequality it suffices to establish its analog for the expected value $\mathbb{E}_\pi(U_t)$ where π is the stationary distribution of the stochastic process U_t. Note that the infinitesimal matrix of the Markov process U_t expressed in the basis $(u_0, u_1, u_2, \ldots)$ is given by

$$Q = \begin{pmatrix} -1 & 1 & 0 & 0 & \cdots \\ r & -(r+1) & 1 & 0 & \cdots \\ r & 0 & -(r+1) & 1 & \\ r & 0 & 0 & \ddots & \ddots \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}$$

By solving $\pi \cdot Q = 0$, we find that

$$\pi = r\left(\frac{1}{r+1}, \left(\frac{1}{r+1}\right)^2, \left(\frac{1}{r+1}\right)^3, \ldots, \left(\frac{1}{r+1}\right)^n, \ldots\right).$$

This implies that

$$\begin{aligned}\mathbb{E}_\nu\left(\bar{X}_t\right) \leq \mathbb{E}_\pi\left(U_t\right) &= r\sum_{n=0}^{\infty} u_n \left(\frac{1}{r+1}\right)^{n+1} \\ &= r\sum_{n=0}^{\infty}(1-(1-\mu)^n)\left(\frac{1}{r+1}\right)^{n+1} \\ &= \frac{r}{r+1}\sum_{n=0}^{\infty}\left(\frac{1}{r+1}\right)^n - \frac{r}{r+1}\sum_{n=0}^{\infty}\left(\frac{1-\mu}{r+1}\right)^n = 1 - \frac{r}{r+\mu}.\end{aligned}$$

The proof of the second inequality is similar. □

Proof of Theorem 8

We first observe that the processes (X_t, Y_t) and $(\bar{X}_t, \bar{Y}_t)$ can be constructed on the same probability space starting from the same initial configuration in such a way that $X_t = \bar{X}_t$ and $Y_t = \bar{Y}_t$ until the hitting time T, which we assume from now on. Let $T_0 = 0$ and, for all $i \geq 1$, let T_i denote the time of the ith jump of the process $\xi_t := \bar{X}_t + \bar{Y}_t$. Since migration events do not change the value of ξ_t, time T_i corresponds to the time of an extinction event at X or an expansion event at Y, therefore we have

$$\begin{aligned} T = T^+ &\quad \text{if and only if there exists } i \geq 0 \\ &\quad \text{such that } T \in (T_i, T_{i+1}) \text{ and } \xi_{T_i} > 2\theta. \\ T = T^- &\quad \text{if and only ifthere exists } i \geq 0 \\ &\quad \text{such that } T \in (T_i, T_{i+1}) \text{ and } \xi_{T_i} < 2\theta. \end{aligned}$$

Let $\epsilon > 0$ small such that $1 - \epsilon > 2\theta$, and consider the events

$$\begin{aligned} D_{i,n}^- &= \{\bar{X}_{T_i} = 0 \text{ and } \bar{Y}_{T_i} \in 2\theta - [n\epsilon, (n+1)\epsilon)\} \\ D_{i,n}^+ &= \{\bar{Y}_{T_i} = 1 \text{ and } \bar{X}_{T_i} \in [n\epsilon, (n+1)\epsilon)\}. \end{aligned}$$

First, since $1 - (n+1)\epsilon > 2\theta - n\epsilon$, migration events between T_i and T_{i+1} displace less individuals on the event $D_{i,n}^-$ than on $D_{i,n}^+$ so

$$P\left(T \in (T_i, T_{i+1}) | D_{i,n}^-\right) \leq P\left(T \in (T_i, T_{i+1}) | D_{i,n}^+\right).$$

Second, note that $v_n = (1-\mu)^n < 2\theta$ if and only if we have

$$n > m_0 := \lfloor \ln(2\theta)/\ln(1-\mu) \rfloor.$$

In particular, if T_i is the time of an extinction event at X then $\bar{Y}_{T_i} < 2\theta$ only if at least m_0 migration events have occurred since the last expansion event at patch Y. This implies that

$$P\left(\bar{Y}_{T_i} < 2\theta\right) \leq \left(\frac{1}{1+r}\right)^{m_0}.$$

Since by symmetry the random variables $\bar{X}_t$ and $1-\bar{Y}_t$ are identically distributed, and $2\theta - \bar{X}_t$ is stochastically smaller than $\bar{Y}_t$, we deduce that

$$P\left(D_{i,n}^-\right) \leq P\left(\bar{Y}_{T_i} < 2\theta\right) P\left(D_{i,n}^+\right) \leq \left(\frac{1}{1+r}\right)^{m_0} P\left(D_{i,n}^+\right).$$

Finally, observing that

$$\{T = T^-\} = \bigcup_{i=0}^{\infty} \bigcup_{n=0}^{\lfloor \epsilon^{-1} \rfloor} \{T \in (T_i, T_{i+1})\} \cap D_{n,i}^-$$

$$\text{and} \quad \{T = T^+\} \supset \bigcup_{i=0}^{\infty} \bigcup_{n=0}^{\lfloor \epsilon^{-1} \rfloor} \{T \in (T_i, T_{i+1})\} \cap D_{n,i}^+$$

we can conclude that

$$\begin{aligned}
P\left(T = T^-\right) &= \sum_{i=0}^{\infty} \sum_{n=0}^{\lfloor \epsilon^{-1} \rfloor} P\left(T \in (T_i, T_{i+1}) | D_{n,i}^-\right) P\left(D_{n,i}^-\right) \\
&\leq \sum_{i=0}^{\infty} \sum_{n=0}^{\lfloor \epsilon^{-1} \rfloor} P\left(T \in (T_i, T_{i+1}) | D_{n,i}^+\right) P\left(D_{n,i}^+\right) \left(P\left(D_{n,i}^-\right)/P\left(D_{n,i}^+\right)\right) \\
&\leq \left(\frac{1}{1+r}\right)^{m_0} \sum_{i=0}^{\infty} \sum_{n=0}^{\lfloor \epsilon^{-1} \rfloor} P\left(T \in (T_i, T_{i+1}) ; D_{n,i}^+\right) \\
&\leq \left(\frac{1}{1+r}\right)^{m_0} P\left(T = T^+\right).
\end{aligned}$$

This completes the proof. □

Acknowledgment The authors would like to thank two anonymous referees for comments that helped to improve the article.

References

Ackleh AS, Allen LJS, Carter J (2007) Establishing a beachhead: a stochastic population model with an Allee effect applied to species invasion. Theor Popul Biol 71:290–300

Adler RF (1993) Migration alone can produce persistence of host-parasitoid models. Am Nat 141:642–650

Amarasekare P (1998a) Allee effects in metapopulation dynamics. Am Nat 152:298–302

Amarasekare P (1998b) Interactions between local dynamics and dispersal: insights from single species models. Theor Popul Biol 53:44–59

Amarasekare P (2000) The geometry of coexistence. Biol J Linn Soc 71:1–31

Berec L, Boukal DS, Berec M (2001) Linking the Allee effect, sexual reproduction, and temperature-dependent sex determination via spatial dynamics. Am Nat 157:217–230

Berec L, Angulo E, Courchamp F (2007) Multiple Allee effects and population management. Trends Ecol Evol 22:185–191

Clark BR, Faeth SH (1997) The consequences of larval aggregation in the butterfly Chlosyne lacinia. Ecol Entomol 22:408–415

Courchamp F, Berec L, Gascoigne J (2009) Allee effects in ecology and conservation. Oxford University Press, Oxford

Dennis B (1989) Allee effects, population growth critical density and the chance of extinction. Natural Resource Modeling 3:481–538

Dennis B (2002) Allee effects in stochastic populations. Oikos 96:389–401

Dercole F, Ferrière R, Rinaldi S (2002) Ecological bistability and evolutionary reversals under asymmetrical competition. Evolution 56:1081–1090

Drake JM (2004) Allee effects and the risk of biological invasion. Risk Anal 24:795–802

Fagan WF, Lewis MA, Neubert MG, van den Driessche P (2002) Invasion theory and biological control. Ecol Lett 5:148–157

Gascoigne JC, Lipcius RN (2004) Allee effects driven by predation. J Appl Ecol 41:801–810

Greene C, Stamps JA (2001) Habitat selection at low population densities. Ecology 82:2091–2100

Guckenheimer J, Holmes P (1983) Nonlinear oscillations, dynamical systems, and bifurcations of vector fields. Springer, New York

Gyllenberg M, Hemminki J, Tammaru T (1999) Allee effects can both conserve and create spatial heterogeneity in population densities. Theor Popul Biol 56:231–242

Harris TE (1972) Nearest neighbor Markov interaction processes on multidimensional lattices. Adv Math 9:66–89

Hopper KR, Roush RT (1993) Mate finding, dispersal, number released, and the success of biological control introductions. Ecol Entomol 18:321–331

Johnson DM, Liebhold AM, Tobin PC, Bjornstad ON (2006) Allee effects and pulsed invasion by the gypsy moth. Nature 444:361–363

Kang Y, Armbruster D (2010) Dispersal effect on a two-patch plant-insect interactions model with Allee-like effects. (Preprint, under review)

Keitt TH, Lewis MA, Holt RD (2001) Allee effects, invasion pinning, and species' borders. Am Nat 157:203–216

Krone SM (1999) The two-stage contact process. Ann Appl Probab 9:331–351

Lande R (1998) Anthropogenic, ecological and genetic factors in extinction and conservation. Res Popul Ecol 40:259–269

Leung B, Drake JM, Lodge DM (2004) Predicting invasions: propagule pressure and the gravity of Allee effects. Ecology 85:1651–1660

Levin SA (1974) Dispersion and population interactions. Am Nat 108:207–228

Liebhold A, Bascompte J (2003) The Allee effect, stochastic dynamics and the eradication of alien species. Ecol Lett 6:133–140

Liebhold A, Tobin PC (2008) Population ecology of insect invasions and their management. Ann Rev Entomol 53:387–408

McCarthy MA (1997) The Allee effect, finding mates and theoretical models. Ecol Model 103:99–102

Petrovskii S, Morozov A, Li B-L (2005) Regimes of biological invasion in a predator-prey system with the Allee effect. Bull Math Biol 67:637–661

Schreiber S (2003) Allee effects, extinctions, and chaotic transients in simple population models. Theor Popul Biol 64:201–209

Shigesada N, Kawasaki K (1997) Introduction. In: Shigesada N, Kawasaki K (eds) Biological invasions: theory and practice. Oxford University Press, New York, pp 1–5

Stephens PA, Sutherland WJ (1999) Consequences of the Allee effect for behaviour, ecology and conservation. Trends Ecol Evol 14:401–405

Taylor CM, Hastings A (2005) Allee effects in biological invasions. Ecol Lett 8:895–908

Tobin PC, Whitmire SL, Johnson DM, Bjrnstad ON, Liebhold AM (2007) Invasion speed is affected by geographical variation in the strength of Allee effects. Ecol Lett 10:36–43

Veit RR, Lewis MA (1996) Dispersal, population growth, and the Allee effect: dynamics of the house finch invasion of eastern North America. Am Nat 148:255–274

Wang MH, Kot M, Neubert MG (2002) Integrodifference equations, Allee effects, and invasions. J Math Biol 44:150–168

Zhou SR, Liu CZ, Wang G (2004) The competitive dynamics of metapopulation subject to the Allee-like effect. Theor Popul Biol 65:29–37