



Extinction risk of a metapopulation under bistable local dynamicsOhad Vilk  and Michael Assaf**Racah Institute of Physics, Hebrew University of Jerusalem, Jerusalem 91904, Israel* (Received 15 August 2019; published 30 January 2020)

We study the extinction risk of a fragmented population residing on a network of patches coupled by migration, where the local patch dynamics includes deterministic bistability. Mixing between patches is shown to dramatically influence the population's viability. We demonstrate that slow migration always increases the population's global extinction risk compared to the isolated case, while at fast migration synchrony between patches minimizes the population's extinction risk. Moreover, we discover a critical migration rate that maximizes the extinction risk of the population, and identify an early-warning signal when approaching this state. Our theoretical results are confirmed via the highly efficient weighted ensemble method. Notably, our theoretical formalism can also be applied to studying switching in gene regulatory networks with multiple transcriptional states.

DOI: [10.1103/PhysRevE.101.012135](https://doi.org/10.1103/PhysRevE.101.012135)**I. INTRODUCTION**

Extinction of a metapopulation—a network of interacting spatially separated populations (patches) of the same species—is of key interest in various scientific disciplines such as ecology, evolutionary biology and genetics [1]. Here a major challenge is finding the optimal interaction strategy among the individual patches in such a fragmented population, that maximizes the metapopulation lifetime. Under certain conditions, it has been found that interactions between the patches decrease the extinction risk of the metapopulation, while in other cases the opposite may occur, i.e., isolation of the individual patches minimizes the population's extinction risk [2,3].

Previous studies of metapopulation dynamics have mostly dealt with the deterministic aspect of the dynamics, using various versions of the Levins model [1,4]. Other approaches, from a more physical point of view, can be found in [5–7]. Several of these models incorporated the so-called Allee effect – a group of effects in population biology which give rise to a negative *per capita* growth rate at small population sizes. For an isolated patch, this effect yields a *colonization threshold*—a critical population density below which extinction occurs deterministically [8]. Thus, under the Allee effect each local patch experiences *bistable* dynamics at the deterministic level. However, in most studies bistability has been modeled only at the metapopulation level [9–13], despite the fact that in many realistic examples of metapopulations, the Allee effect has been shown to play a crucial role in the dynamics of the population [1,14–16].

Incorporating the Allee effect at the individual patch level can greatly impact the population's extinction risk, and thus affects population management and preservation [17,18]. Indeed, the Allee effect strongly influences both extinction and colonization of individual patches, and thus it is vital to take

it into account when dealing with metapopulation extinction [1,19]. Yet, while recent studies provided an explicit calculation of the mean time to extinction (MTE) in such systems [19–23], none of these works has conducted a systematic study of the metapopulation's extinction risk, while incorporating bistable dynamics at the individual patch level.

In this paper, by coupling the local demographic noise to stochastic migration between patches, we investigate the metapopulation behavior when the local birth-death dynamics on each patch exhibit the Allee effect in the form of bistable dynamics. At slow migration the system displays multiple routes to extinction, and we find that counterintuitively weak migration always reduces the metapopulation's stability. This indicates that isolation is preferable to weak mixing. In particular, our result challenges the long-standing “rescue effect,” which states that the rescue of local patches necessarily increases the metapopulation viability [1,21]. At fast migration, we show that migration gives rise to synchronization between patches and the extinction risk is minimized when the typical flux across patches is comparable. In extreme cases where such synchrony cannot be achieved, one observes *source-sink* dynamics, where one patch becomes a sink to the others. Our theoretical results are tested against highly efficient numerical simulations based on the weighted-ensemble (WE) method. Importantly, our analysis provides the exact conditions for which mixing (at some migration rate) or complete isolation is optimal for the population's viability, which may have important consequences on real-life populations.

The paper is organized as follows. In Sec. II we analyze the deterministic dynamics of interconnected patches. In Sec. III we incorporate demographic noise and analyze the corresponding master equation using the Wentzel-Kramers-Brillouin (WKB) method. In Sec. III A we show that for slow migration there exist multiple routes to extinction, and compute the global MTE, as well as local extinction and colonization rates. Here we also identify a critical migration rate that maximizes the global extinction risk. In Sec. III B we compute the MTE for fast migration, up to subleading-order

*michael.assaf@mail.huji.ac.il

corrections, where we also find an optimal migration rate that maximizes the population's lifetime. In Sec. IV we describe the WE simulation, while in Sec. V we summarize our results and discuss additional applications of our model in other scientific disciplines.

II. MODEL AND DETERMINISTIC DYNAMICS

We consider M patches, where migration between patch i and j occurs at a rate μ_{ij} . To locally account for the Allee effect on each patch $i = 1, \dots, M$, we choose a simple birth-death process, $2A \leftrightarrow 3A$ and $A \rightarrow 0$, that gives rise to bistable dynamics [22]. As a result, on each patch we have a single-step birth-death process [19]:

$$n_i \xrightarrow{B_{n_i}} n_i + 1, \quad n_i \xrightarrow{D_{n_i}} n_i - 1, \quad n_i \xrightarrow{\mu_{ij} n_i} n_j, \\ B_{n_i} = \frac{2n_i(n_i - 1)}{N_i(1 - \delta_i^2)}, \quad D_{n_i} = n_i + \frac{n_i(n_i - 1)(n_i - 2)}{N_i^2(1 - \delta_i^2)}, \quad (1)$$

where n_i denotes the population size on patch i , and the reaction $n_i \rightarrow n_j$ denotes migration from patch i to patch j . In addition, $0 < \delta_i < 1$ determines the distance between the local colonized state and colonization threshold, while $N_i \gg 1$ is the local carrying capacity; see below. Denoting $\kappa_i = N_i/N_1$, $\alpha_{ij} = \mu_{ij}/\mu$, and $x_i = n_i/N_1$ as the population density on patch i , the deterministic rate equation for the average population density \bar{x}_i , reads

$$\dot{\bar{x}}_i = b_i(\bar{x}_i) - d_i(\bar{x}_i) + \mu \sum_{j=1}^M (\alpha_{ji} \bar{x}_j - \alpha_{ij} \bar{x}_i). \quad (2)$$

Here μ is the characteristic migration rate, while $b_i(x_i) = B_{n_i}/N_1$ and $d_i(x_i) = D_{n_i}/N_1$, such that

$$b_i(x_i) = \frac{2x_i^2}{\kappa_i(1 - \delta_i^2)}, \quad d_i(x_i) = x_i + \frac{x_i^3}{\kappa_i^2(1 - \delta_i^2)}, \quad (3)$$

where $O(N^{-1})$ terms were neglected.

While most of the analysis below can be generalized for M patches; see Appendix B, here we focus on the dynamics of two connected patches, which capture most of the interesting features in this problem. Thus, henceforth we have $M = 2$ and denote $\alpha_{12} \equiv 1$ and $\alpha_{21} \equiv \alpha$, where $\alpha = O(1)$ is the ratio between the migration rates, while $\kappa_1 = 1$ and $\kappa_2 \equiv \kappa$ is the carrying capacities ratio. We also denote $N \equiv N_1$ such that $N_2 = \kappa N$. For two patches, rate equations (2) become

$$\dot{\bar{x}}_1 = b_1(\bar{x}_1) - d_1(\bar{x}_1) + \alpha \mu \bar{x}_2 - \mu \bar{x}_1, \\ \dot{\bar{x}}_2 = b_2(\bar{x}_2) - d_2(\bar{x}_2) + \mu \bar{x}_1 - \alpha \mu \bar{x}_2. \quad (4)$$

At zero migration, $\mu = 0$, each patch has three fixed points (FPs). Putting $\dot{\bar{x}}_1 = \dot{\bar{x}}_2 = 0$ in Eqs. (4) we find

$$x_{i,0}^{(0)} = 0, \quad x_{i,\pm}^{(0)} = \kappa_i(1 \pm \delta_i), \quad (5)$$

where $x_{i,0}^{(0)}$ and $x_{i,+}^{(0)}$ are stable FPs, respectively corresponding to the extinction and colonized states, and $x_{i,-}^{(0)}$ is unstable and corresponds to the colonization threshold [8]. Thus, there are nine FPs for zero migration. Note that the relaxation time in the vicinity of the colonized state, which equals $(1 - \delta_i)/(2\delta_i)$, determines the typical timescale of the local patch dynamics, where in general it is $O(1)$; see below.

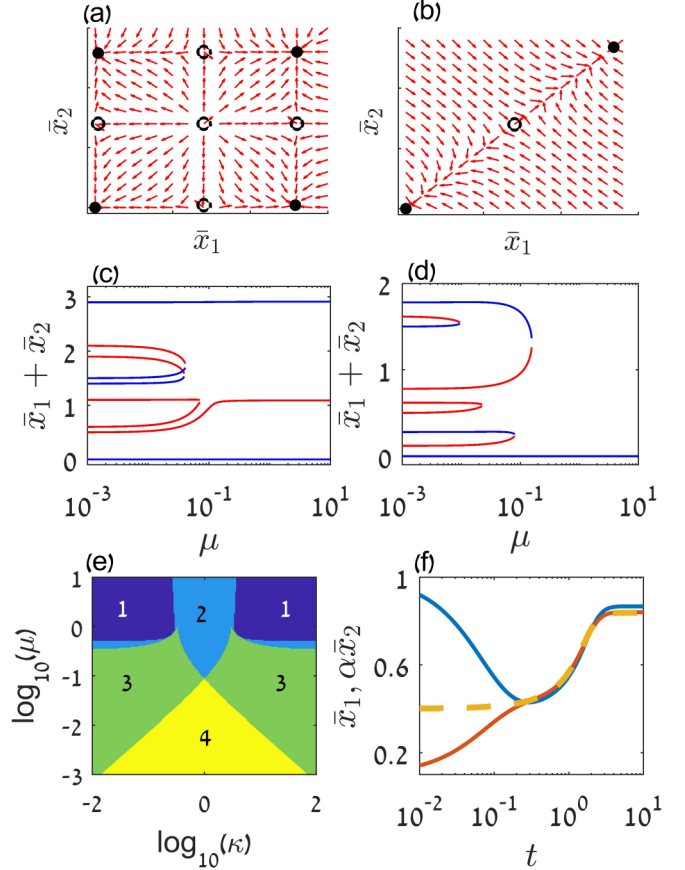


FIG. 1. (a) and (b) Dynamical trajectories of Eqs. (4) in the case of two patches, with $\kappa = 1$, $\alpha = 1$, $\delta_1 = \delta_2 = 0.3$, and (a) $\mu = 0.01$, (b) $\mu = 1$. Stable and unstable FPs are denoted by full and open circles, respectively. (c) and (d) Bifurcation diagram of Eqs. (4), as a function of μ , with $\alpha = 1$, $\delta_1 = 0.5$, $\delta_2 = 0.4$, and (c) $\kappa = 1$, (d) $\kappa = 0.2$. Here the blue and red lines correspond to stable and unstable FPs, respectively. (e) Number of stable FPs as function of both κ and μ , for $\alpha = 1$, $\delta_1 = \delta_2 = 0.5$. (f) \bar{x}_1 and $\alpha \bar{x}_2$ as a function of time by numerically solving Eqs. (4) (upper blue and lower red lines), compared to the numerical solution of Eq. (11) for ξ (dashed line). Here $\kappa = 1$, $\delta_1 = 0.5$, $\delta_2 = 0.6$, $\mu = 10$, $\alpha = 0.5$.

A. The case of slow migration

For slow migration, $\mu \ll 1$, i.e., when the typical time scale of migration is slow compared to that of the local patch, Eqs. (4) give rise to a maximum of nine FPs, four of which are stable; see Figs. 1(a) and 1(c)–1(e). Here, the FPs are shifted by $O(\mu)$ compared to the case where the patches are isolated. To find the shift in the FPs, we look for the solution as $\mathbf{x}_s \equiv [x_{s,1}, x_{s,2}] = (x_{1,s_1}^{(0)} + \mu \eta_{1,s}, x_{2,s_2}^{(0)} + \mu \eta_{2,s})$ [24] for $s = (s_1, s_2)$ with $s_i = \{0, +, -\}$ representing the possible states. Here $x_{i,s_i}^{(0)}$ are given by Eqs. (5) and $\eta_{i,s}$ are the yet unknown shifts. Substituting this solution into Eqs. (4), putting $\dot{\bar{x}}_1 = \dot{\bar{x}}_2 = 0$ and keeping terms up to $O(\mu)$, we find $\eta_{i,s}$, which yields the FPs up to $O(\mu)$:

$$(\mathbf{x}_s)_1 = x_{1,s_1}^{(0)} + \mu t_{1,s_1}^{\text{relax}} (\alpha x_{2,s_2}^{(0)} - x_{1,s_1}^{(0)}), \\ (\mathbf{x}_s)_2 = x_{2,s_2}^{(0)} + \mu t_{2,s_2}^{\text{relax}} (x_{1,s_1}^{(0)} - \alpha x_{2,s_2}^{(0)}). \quad (6)$$

Here we have defined $t_{i,s_i}^{\text{relax}} = [b'_i(x_{i,s_i}^{(0)}) - d'_i(x_{i,s_i}^{(0)})]^{-1}$, as the relaxation time to the FP $x_{i,s_i}^{(0)}$, at the level of the isolated patch. This result [Eq. (6)] can be intuitively understood as follows: in the leading order the average population density in each patch is determined by the outgoing flux from itself and incoming flux from the second patch. The magnitude of the correction depends on the migration rate as well as on the relaxation time to the relevant FP. Note that since $s = (s_1, s_2)$ can receive nine different values, Eq. (6) represents nine FPs. It can be shown via linear stability analysis that four of these FPs are stable while five of them are unstable. The former correspond to scenarios where either both patches are colonized, one patch is colonized and the other is close to extinction, and both patches are extinct. This can be observed in the different panels of Fig. 1. In Fig. 1(a) we plot dynamical trajectories of Eqs. (4) for slow ($\mu = 0.01$) migration. Furthermore, in Figs. 1(c)–1(e) one can see a total of four stable and five unstable FPs at small μ .

For concreteness we present two examples of the FPs given by Eqs. (6). When patch 1 is colonized and patch 2 is close to extinction, the stable FP is given by

$$\begin{aligned} \mathbf{x}_{+,0} &= [(\mathbf{x}_{+,0})_1, (\mathbf{x}_{+,0})_2] \\ &= [1 + \delta_1 - \mu(1 - \delta_1^2)/(2\delta_1), \mu(1 + \delta_1)], \end{aligned} \quad (7)$$

while in the opposite case where patch 2 is colonized and patch 1 is close to extinction, the stable FP is given by

$$\begin{aligned} \mathbf{x}_{0,+} &= [(\mathbf{x}_{0,+})_1, (\mathbf{x}_{0,+})_2] \\ &= [\mu\alpha\kappa(1 + \delta_2), \kappa(1 + \delta_2) - \mu\alpha\kappa(1 - \delta_2^2)/(2\delta_2)]. \end{aligned} \quad (8)$$

Here we have used the fact that $t_{i,+}^{\text{relax}} = (1 - \delta_i)/(2\delta_i)$ and $t_{i,0}^{\text{relax}} = 1$; see the definition following Eq. (6). In Eqs. (7) and (8) it is evident that the correction to the colonized patch due to slow migration is negative, as deterministically, the incoming flux from the patch close to extinction contributes only $O(\mu^2)$ terms.

B. The case of fast migration

For fast migration, $\mu \gg 1$, Eqs. (4) give rise to a maximum of three FPs, two of which are stable; see Figs. 1(b) and 1(e). Indeed, as μ increases the overall number of FPs decreases via multiple bifurcations; see Figs. 1(c) and 1(d).

To find the FPs in the leading order in μ , we put $\dot{x}_1 = \dot{x}_2 = 0$ in Eqs. (4), and keep only terms proportional to μ . This yields $\bar{x}_2 = \bar{x}_1/\alpha$, meaning that the patches are synchronized in the leading order. Summing Eqs. (4) and substituting $\bar{x}_1 = \xi$, $\bar{x}_2 = \xi/\alpha$ we find the following FPs in terms of ξ :

$$\xi_0 = 0, \quad \xi_{\pm} = \tilde{\kappa}(1 \pm \tilde{\delta}). \quad (9)$$

These FPs coincide with those in Eqs. (5), upon replacing δ and κ by their effective counterparts:

$$\begin{aligned} \tilde{\kappa} &= \frac{\alpha\kappa[\alpha^2\kappa(1 - \delta_2^2) + 1 - \delta_1^2]}{\alpha^3\kappa^2(1 - \delta_2^2) + 1 - \delta_1^2}, \\ \tilde{\delta} &= \left[1 - \frac{\alpha(\alpha + 1)\kappa(1 - \delta_1^2)(1 - \delta_2^2)}{\tilde{\kappa}[\alpha^2\kappa(1 - \delta_2^2) + 1 - \delta_1^2]} \right]^{1/2}. \end{aligned} \quad (10)$$

Here $\tilde{\kappa}$ and $\tilde{\delta}$ are the effective carrying capacity and threshold parameters, respectively. Note that $\tilde{\kappa} > 0$, since $\delta_i < 1$ for $i = 1, 2$. Furthermore, we have assumed $0 < \tilde{\delta} < 1$; otherwise $\tilde{\delta}$ is imaginary and the extinction state is the only stable FP; see below.

In addition to finding the FPs one can also study the relaxation dynamics by using the definition of ξ and summing Eqs. (4). As a result, the effective rate equation for ξ becomes [19]

$$\dot{\xi} = -\xi + \frac{2\xi^2}{\tilde{\kappa}(1 - \tilde{\delta}^2)} - \frac{\xi^3}{\tilde{\kappa}^2(1 - \tilde{\delta}^2)}. \quad (11)$$

In Fig. 1(f) we demonstrate the synchrony between \bar{x}_1 and $\alpha\bar{x}_2$ by comparing numerical solutions of Eqs. (4) and Eq. (11). As can be observed in this panel, synchronization between the two patches is achieved at times $t \gtrsim O(\mu^{-1})$, making Eq. (11) valid at times $t \gg O(\mu^{-1})$.

A simple example demonstrating bistability, i.e., two stable FPs in Eq. (11), is when $\delta_1 = \delta_2 = \delta$ and $\kappa = \alpha^{-1}$. This yields $\tilde{\kappa} = 1$ and $\tilde{\delta} = \delta$, and thus, the FPs in Eqs. (9) become $\xi_{\pm} = 1 \pm \delta$. This is not a trivial result as it predicts that when α and κ counter each other such that the typical flux between patches is approximately equal, the resulting dynamics mimic those of the original patches. In this case, the system is said to be well mixed [19].

A transition between bistability and monostability (a single FP at extinction) can occur, e.g., when the carrying capacities are very different, $\kappa \ll 1$, where $\alpha = O(1)$. In this case, we find in the leading order of κ , $\tilde{\kappa} \simeq \kappa\alpha$ and $\tilde{\delta} \simeq [\delta_2^2 - \alpha(1 - \delta_2^2)]^{1/2} \equiv \Delta_2$. Since these depend only on the parameters of patch 2, which has a much smaller carrying capacity, this entails that the smaller patch dictates the deterministic size of the metapopulation. Here bistability is obtained as long as $\delta_2 \geq [\alpha/(1 + \alpha)]^{1/2}$ [25].

The different scenarios of bistability and monostability at large μ are demonstrated in Figs. 1(c)–1(e). In Figs. 1(c) and 1(d), by numerically solving Eqs. (4) for the entire range of μ , we demonstrate the multiple bifurcations that occur. Starting from nine FPs at small μ , as μ is increased the system ends up with either one or three FPs. In Fig. 1(e) we map the number of stable FPs as a function of both μ and κ , displaying a reduction in the number of FPs as μ increases, and as κ diverges from α^{-1} . Here, as μ is increased, the number of stable FPs tends either to one or two depending on the value of κ (see discussion above) and the other parameters.

Finally, we can compute the subleading $O(\mu^{-1})$ corrections to the FPs [Eqs. (9)] in the limit of $\mu \gg 1$. While we do not give the explicit expressions here, these will be used to compute the subleading-order corrections to the MTE at fast migration; see Sec. III B.

III. STOCHASTIC FORMULATION

To account for local demographic noise and the stochastic migration across patches, we write down the master equation describing the evolution of \mathbb{P}_{n_1, n_2} – the probability of finding n_1 and n_2 individuals in patch 1 and 2, respectively,

at time t . Using Eqs. (1), the master equation yields

$$\dot{\mathbb{P}}_{n_1, n_2} = \left\{ \sum_{i=1}^2 [(\mathbb{E}_{n_i}^{-1} - 1)B_{n_i} + (\mathbb{E}_{n_i}^1 - 1)D_{n_i}] + [(\mathbb{E}_{n_1}^1 \mathbb{E}_{n_2}^{-1} - 1)n_1 + (\mathbb{E}_{n_1}^{-1} \mathbb{E}_{n_2}^1 - 1)\alpha n_2] \mu \right\} \mathbb{P}_{n_1, n_2}, \quad (12)$$

where $\mathbb{E}_{n_i}^{\pm 1} f(n) = f(n \pm 1)$. In the absence of external flux into either of the patches, starting from any initial condition, the system ultimately undergoes extinction, where $\mathbb{P}_{0,0}$ grows in time while all other probabilities decay. Yet, in the limit of large carrying capacities, the decay rate turns out to be exponentially small, see below, and one can use the metastable ansatz $\mathbb{P}_{n_1, n_2} = \pi_{n_1, n_2} \exp(-t/\tau)$, where π_{n_1, n_2} is the quasistationary distribution and τ is the MTE. The latter can be found by employing the WKB ansatz $\pi_{n_1, n_2} \sim \exp[-NS(x_1, x_2)]$, where S is the action function [26], arriving at a stationary Hamilton-Jacobi equation, $H = 0$ [26–31], with Hamiltonian

$$H(x_1, p_1, x_2, p_2) = \sum_{i=1}^2 (e^{p_i} - 1)[b_i(x_i) - e^{-p_i} d_i(x_i)] + x_1 \mu (e^{p_2 - p_1} - 1) + x_2 \mu \alpha (e^{p_1 - p_2} - 1), \quad (13)$$

where $p_i = \partial_{x_i} S$ are the conjugate momenta, and we have used Eqs. (3). Hamiltonian (13) yields a set of four Hamilton equations, which can be solved numerically for any set of parameters, yielding the MTE [32,33]. Analytical progress can be made in two limiting cases: slow and fast migration.

A. The case of slow migration

For slow migration, $\mu \ll 1$, in general there are four stable FPs at the deterministic level. However, when accounting for demographic noise, these FPs become metastable states, which means that the system can stochastically switch between any pair of them. Importantly, the presence of multiple metastable states gives rise to multiple extinction routes. To find the MTE of the metapopulation, we apply a similar method to Ref. [34], and define $\mathcal{P}_1 \equiv \mathcal{P}(\{x_{+,+}\})$, $\mathcal{P}_2 \equiv \mathcal{P}(\{x_{+,0}\})$, $\mathcal{P}_3 \equiv \mathcal{P}(\{x_{0,+}\})$, $\mathcal{P}_4 \equiv \mathcal{P}(\{x_{0,0}\})$, as the probabilities to be at the basins of attraction of the FPs where both patches are colonized (FP1), patch 1 is colonized and patch 2 is extinct (FP2), patch 2 is colonized and patch 1 is extinct (FP3), and both patches are extinct (FP4); see Fig. 2(a) for an illustration. Assuming the transition rates r_{ij} between \mathcal{P}_i and \mathcal{P}_j are known (to be calculated below), the probabilities \mathcal{P}_i ($i = 1, 2, 3, 4$) satisfy

$$\dot{\mathcal{P}}_i(t) = \sum_{j \neq i} r_{ji} \mathcal{P}_j(t) - r_{ij} \mathcal{P}_i(t), \quad (14)$$

where the MTE is given by $\tau = \int_0^\infty t \dot{\mathcal{P}}_4 dt$ [34,35]. While these equations can be solved for any r_{ij} (see Appendix A) it turns out that the solution drastically simplifies for any reasonable choice of parameters, and satisfies

$$\tau \simeq \min \left\{ \max[r_{12}^{-1}, r_{24}^{-1}, r_{21}/(r_{12}r_{24})], \max[r_{13}^{-1}, r_{34}^{-1}, r_{31}/(r_{13}r_{34})] \right\}; \quad (15)$$

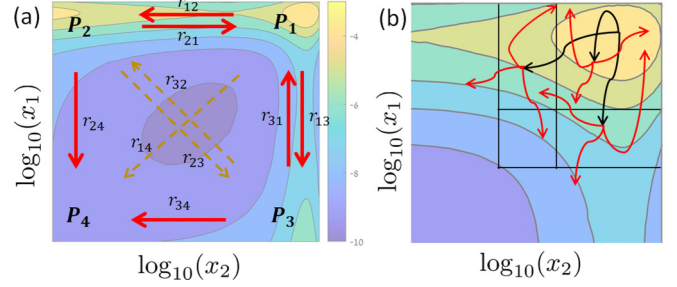


FIG. 2. Phase space and WE simulations. (a) Transitions between basins of attraction of FPs 1–4 (see Sec. III) drawn over the quasistationary distribution obtained by the WE simulation (see Sec. IV). (b) Illustration depicting the two steps we repeat in a WE simulation, propagation and resampling (see Sec. IV). Parameters in both panels are $N = 100$, $\kappa = 1$, $d_1 = 0.5$, $d_2 = 0.4$, $\mu = 10^{-3}$.

see Appendix A for details. Here we have assumed that switches, which involve synchronous transitions of *both* patches, occur at an exponentially slower rate than other transitions; i.e., rates r_{14} , r_{23} , r_{32} , are negligible compared to all other rates; see Appendix A. The outer minimum in Eq. (15) chooses the extinction route with the overall minimal cost, while the inner maximum determines the cost of the chosen trajectory.

1. Transition rates calculation

We now compute the rates r_{ij} in the limit of $\mu \ll 1$, while taking into account the fact that in this limit, extinction occurs in a *serial* manner with an overwhelming probability; see above. Let us start by computing the extinction rates, and without loss of generality let us consider the extinction of patch 1 while patch 2 remains colonized. That is, we assume that x_2 and p_2 fluctuate around their colonized FP with demographic noise of order $(\kappa N)^{-1/2}$, and additional migrational noise of order μ , where both μ and $(\kappa N)^{-1/2}$ are small and uncorrelated. To this end, we substitute $x_2 = x_{2,+}^{(0)}[1 + O(\mu) + O(N^{-1/2})]$ and $p_2 = O(\mu) + O(N^{-1/2})$, into Hamiltonian (13), and keep terms up to first order in $\mu \ll 1$ and zeroth order in $N \gg 1$. After some algebra, this yields an effective Hamiltonian that accounts for the transition FP1 \rightarrow FP3, i.e., the extinction of patch 1, while experiencing patch 2 as constant external flux at its colonized state [see Fig. 2(a)]:

$$H_{\text{eff}}^{\text{slow}}(x_1, p_1) = (e^{p_1} - 1)[b_1(x_1) - e^{-p_1} d_1(x_1)] + \mu(e^{-p_1} - 1)x_1 + \alpha \mu \kappa (1 + \delta_2)(e^{p_1} - 1). \quad (16)$$

From this Hamiltonian we can compute the optimal path from FP1 to FP3—a heteroclinic trajectory the system takes with an overwhelming probability, connecting FP1 and FP3 [26]. This path can be found by equating $H_{\text{eff}}^{\text{slow}}(x_1, p_1) = 0$ which yields

$$p_1(x_1) = \ln \left[\frac{d_1(x_1) + \mu x_1}{b_1(x_1) + \mu \alpha \kappa (1 + \delta_2)} \right], \quad (17)$$

which, for $\mu \ll 1$, can be approximated as

$$p_1(x_1) \simeq \ln \left[\frac{d_1(x_1)}{b_1(x_1)} \right] + \mu \left[\frac{x_1}{d_1(x_1)} - \frac{\alpha \kappa (1 + \delta_2)}{b_1(x_1)} \right]. \quad (18)$$

Therefore, the action satisfies

$$S_{13} = \sum_{k=1}^2 \int_{(x_{+,+})_k}^{(x_{-,+})_k} p_k(x_k) dx_k \simeq \int_{(x_{+,+})_1}^{(x_{-,+})_1} p_1(x_1) dx_1, \quad (19)$$

where the integral limits are given by Eqs. (6), and the integral over $p_2(x_2)$ was ignored as it contributes only $O(\mu^2)$ terms. Performing the integral in Eq. (19) with $p_1(x_1)$ given by (18), we find the action S_{13} up to first order in μ . Note that, computing the actions S_{12} , S_{24} , and S_{34} corresponding to the transitions between FP1 and FP2, FP2 and FP4, and FP3 and FP4, respectively [see Fig. 2(a)], can be done in a similar manner. These calculations yield the following extinction actions:

$$\begin{aligned} S_{24} &= S_0(\delta_1)(1 + \mu/2) - \mu\delta_1, \\ S_{13} &= S_{24} + \mu\alpha\kappa\delta_1(1 + \delta_2), \\ S_{34} &= \kappa S_0(\delta_2)(1 + \mu\alpha/2) - \mu\alpha\kappa\delta_2, \\ S_{12} &= S_{34} + \mu\delta_2(1 + \delta_1), \end{aligned} \quad (20)$$

where

$$S_0(\delta_i) = 2 \left[\delta_i - (1 - \delta_i^2)^{1/2} \arcsin(\delta_i) \right] \quad (21)$$

is the action of isolated patch i . Having found these actions, in the leading order in $N \gg 1$, the extinction rates are given by $r_{ij} = e^{-NS_{ij}}$ for $ij = \{12, 13, 34, 24\}$.

Finding the colonization rates r_{21} and r_{31} , can be done in a similar manner by integrating over optimal path (17) between the corresponding FPs. For example, the colonization rate of patch 1 while patch 2 is colonized [transition FP3 \rightarrow FP1; see Fig. 2(a)] is found by integrating between $(x_{0,+})_1$ and $(x_{-,+})_1$, given by Eqs. (6), and approximating the result up to $O(\mu)$ [36]. The transition FP2 \rightarrow FP1 can be found in a similar manner. This yields the colonization actions

$$\begin{aligned} S_{31} &= \left\{ \delta_1 + 2(1 - \delta_1^2)^{1/2} \arcsin \left[\left(\frac{1 - \delta_1}{2} \right)^{1/2} \right] - 1 \right\} \\ &\quad - \mu^{1/2} \frac{\pi [\kappa\alpha(1 + \delta_2)(1 - \delta_1^2)]^{1/2}}{2^{1/2}} \\ &\quad + \mu \left\{ (1 - \delta_1^2)^{1/2} \arcsin \left[\left(\frac{1 - \delta_1}{2} \right)^{1/2} \right] + \frac{\kappa\alpha(1 + \delta_2)(\delta_1 + 3)}{2} \right\} \end{aligned} \quad (22)$$

and

$$\begin{aligned} S_{21} &= \kappa \left\{ \delta_2 + 2(1 - \delta_2^2)^{1/2} \arcsin \left[\left(\frac{1 - \delta_2}{2} \right)^{1/2} \right] - 1 \right\} \\ &\quad - \mu^{1/2} \frac{\pi [(1 + \delta_1)(1 - \delta_2^2)\kappa]^{1/2}}{2^{1/2}} \\ &\quad + \mu \left\{ \alpha(1 - \delta_2^2)^{1/2} \kappa \arcsin \left[\left(\frac{1 - \delta_2}{2} \right)^{1/2} \right] + \frac{(1 + \delta_1)(\delta_2 + 3)}{2} \right\}. \end{aligned} \quad (23)$$

The colonization rates are thus given by $r_{21} = \exp(-NS_{21})$ and $r_{31} = \exp(-NS_{31})$.

Notably, Eqs. (15), (20), (22), and (23) show the advantage of our formulation compared to early metapopulation models, where deterministic bistability was accounted for by averaging over multiple FPs in the system [1, 10, 13, 23]. Our approach fully accounts for multiple metastable states, and may be used to assess the validity of such averaging. Our

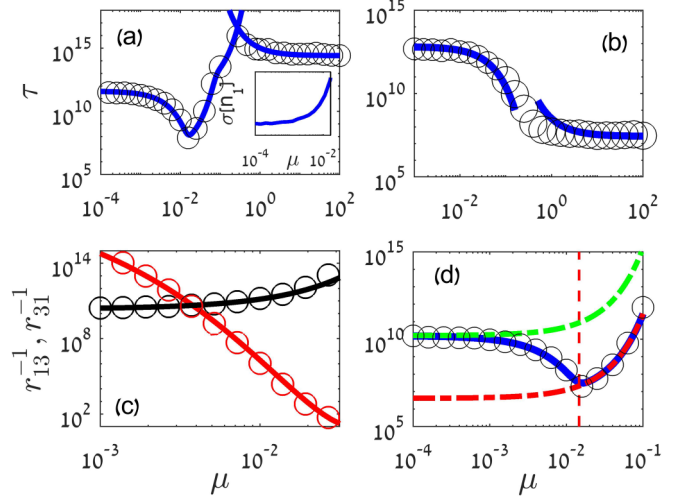


FIG. 3. (a) and (b) MTE as a function of μ : theory (solid line) given by Eq. (15) with Eqs. (20) for slow migration, and Eq. (32) for fast migration, compared to WE simulations (circles). Here (a) $N = 2300$, $\kappa = 0.87$, $\delta_1 = 0.25$, $\delta_2 = 0.21$, $\alpha = 1$; (b) $N = 400$, $\kappa = 0.25$, $\delta_1 = 0.45$, $\delta_2 = 0.7$, $\alpha = 1$. Inset of (a) shows the sharp increase in the standard deviation of patch 1 around its stable FP (normalized by the mean size of both patches) when approaching $\mu_{\text{crit}} \simeq 1.5 \times 10^{-2}$. (c) Comparison of the extinction (increasing black line) and colonization (decreasing red line) rates, for $N = 250$, $\kappa = 0.5$, $\delta_1 = 0.62$, $\delta_2 = 0.6$, $\alpha = 1$. (d) Shown are the MTE (solid line) [Eq. (15)], the local extinction rates FP1 \rightarrow FP3 (upper green dashed line) and FP1 \rightarrow FP2 (lower red dashed line) [see Eqs. (20)], and WE simulations (circles) as a function of μ ; the dashed vertical line denotes μ_{crit} [Eq. (24)]. Here $N = 2000$, $\kappa = 1$, $\delta_1 = 0.25$, $\delta_2 = 0.21$, and $\alpha = 1$.

analysis also allows for calculating these transition rates from first principles.

In Figs. 3(a) and 3(b) we compare, for two different parameter sets, our analytical result for τ versus μ [Eq. (15)], with highly efficient WE simulations [37–39]; see Figs. 2(b) and Sec. IV for details. In Fig. 3(c) we compare extinction and colonization rates of a local patch (when the second patch is colonized), which are shown to strongly depend on migration, in contrast to some early metapopulation models [1]. Importantly, for sufficiently slow migration τ decreases as μ increases; see Figs. 3(a) and 3(b). That is, as μ is increased from zero, the metapopulation's extinction risk increases, compared to the isolated case.

To understand why this occurs, let us analyze the competing terms in Eq. (15). For $\mu \rightarrow 0$, the colonization rates r_{21} , r_{31} vanish, and thus, at sufficiently slow migration, these rates can be neglected. Moreover, while at $\mu = 0$ we have $r_{13} = r_{24}$ and $r_{12} = r_{34}$, see Eqs. (20), as μ is increased, r_{24} and r_{34} respectively increase at a faster rate than r_{13} and r_{12} (which do not necessarily increase at all). Thus, as μ increases, the minimum in Eq. (15) necessarily chooses the terms r_{24}^{-1} over r_{13}^{-1} and r_{34}^{-1} over r_{12}^{-1} ; i.e., the MTE is determined by the maximum of r_{24}^{-1} and r_{34}^{-1} , both of which decrease as μ is increased.

Why do r_{24}^{-1} and r_{34}^{-1} decrease? Along the transitions FP2 \rightarrow FP4 and FP3 \rightarrow FP4 there is one patch that is

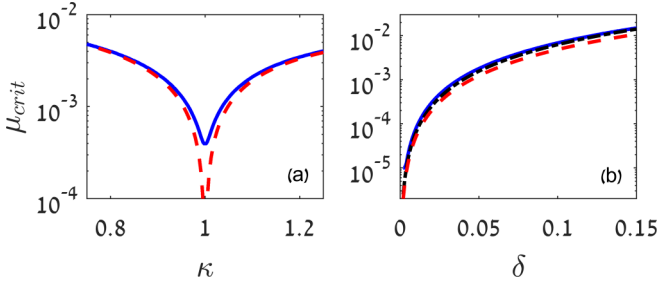


FIG. 4. (a) μ_{crit} versus κ as predicted by the exact form of τ given by Eq. (A7) (solid line) and as given by Eq. (24) (dashed line). Parameters are $\delta_1 = \delta_2 = 0.2$, $\alpha = 1$. (b) μ_{crit} as a function of $\delta \equiv \delta_1 = 2\delta_2$. The solid line is the exact result as obtained from Eq. (A7), the dashed line is Eq. (24), and the dashed-dotted line is the bifurcation result [Eq. (25)]. Parameters are $\kappa = 1/\alpha = 1/2$.

colonized and another, close to extinction. As μ grows, the colonized patch sends individuals to the patch close to extinction, while the back flux is negligible. Thus, whereas the flux from the colonized patch increases its extinction risk (due to loss of individuals), it cannot rescue the other patch from extinction, as its population size is below the colonization threshold. Hence, weak migration increases the metapopulation's global extinction risk compared to the isolated case, which is a direct consequence of the Allee effect and the existence of a colonization threshold; for local logistic dynamics the opposite is observed [20].

The decrease of $\tau(\mu)$ at small μ can give rise to another fascinating phenomenon: the existence of a global minimum of $\tau(\mu)$ at a critical (and finite) migration rate μ_{crit} , that maximizes the extinction risk of the metapopulation; see Sec. III A 2. In Fig. 3(a) we observe such a global minimum, while in Fig. 3(b), the minimum is obtained only at $\mu \rightarrow \infty$. Empirically, approaching the minimum of τ is accompanied by a sharp increase in the population's variance of the colonized patch (in FP2/FP3) as μ approaches μ_{crit} ; see inset of Fig. 3(a). This increase can serve as an early warning signal for stability deterioration of the entire metapopulation [40].

Notably, our analysis invalidates a long standing claim, that the rescue of local patches necessarily increases the metapopulation's viability [1,21], when the local dynamics include the Allee effect. Indeed, for small μ , as μ increases a single patch can experience a decrease in its local extinction risk ("rescue effect") even though the global extinction risk increases. This behavior is apparent by comparing the MTE of the global metapopulation with the MTE of a single patch while the other patch is colonized; see Fig. 3(d). Here, while each patch is shown to locally profit from increasing migration, the MTE of the global metapopulation decreases with increasing μ .

2. Critical migration rate

As specified above, in some cases there exists a critical migration rate, μ_{crit} , for which the metapopulation's extinction risk is maximized. In general, μ_{crit} can be computed numerically using the exact form of the MTE [Eq. (A7) in Appendix A]; see Fig. 4. For simplicity, here we find μ_{crit} in the case of negligible colonization rates [41]. In this case, the maximum functions in Eq. (15) choose between r_{24}^{-1} and r_{12}^{-1}

or between r_{34}^{-1} and r_{13}^{-1} . As a result, the critical migration rate is given by equating $S_{24} = S_{12}$ or $S_{34} = S_{13}$:

$$\mu_{\text{crit}} = \frac{S_0(\delta_1) - \kappa S_0(\delta_2)}{\delta_1 - S_0(\delta_1)/2 + C - \alpha\kappa[\delta_2 - S_0(\delta_2)/2]}. \quad (24)$$

Here C is determined by the relative stability of the patches. For $S_0(\delta_1) > \kappa S_0(\delta_2)$, $C = \delta_2(\delta_1 + 1)$. In this case μ_{crit} is a global minimum as long as the correction in S_{12} is positive [see Eqs. (20)], i.e., $\delta_2(\delta_1 + 1) > \alpha\kappa\delta_2 - S_0(\delta_2)/2$. On the other hand, for $S_0(\delta_1) < \kappa S_0(\delta_2)$, $C = \alpha\kappa\delta_1(\delta_2 + 1)$. In this case, μ_{crit} is a global minimum as long as the correction in S_{13} is positive, i.e., $\alpha\kappa\delta_1(\delta_2 + 1) > \delta_1 - S_0(\delta_1)/2$.

In Fig. 4 we plot μ_{crit} as a function of κ and δ . In Fig. 4(a) we compare Eq. (24) with the critical migration rate obtained numerically by finding the minimum of the exact form of τ , given by Eq. (A7) in Appendix A. We find that Eq. (24) is a good approximation as long as κ is not too close to 1; otherwise, the assumption that the transition rates exponentially differ from each other breaks down, which invalidates Eq. (15) and Eq. (24).

Equation (24) drastically simplifies close to the bifurcation limit where $\delta_1, \delta_2 \ll 1$. To remind the reader, at the bifurcation limit the colonized state and colonization threshold merge [31]. Taking $\delta_1 \equiv \delta$ and $\delta_2 = u\delta$, where $u = O(1)$, Eq. (24) simplifies to

$$\mu_{\text{crit}} \simeq \max\{(2/3)\delta^2(1 - \kappa u^3), [2/(3u)]\delta^2(\kappa u^3 - 1)\}, \quad (25)$$

which agrees well with Eq. (24) for $\delta \ll 1$; see Fig. 4(b).

B. The case of fast migration

For fast migration, $\mu \gg 1$, only FPs 1 and 4 remain; see Figs. 1(b) and 1(c). Thus, extinction can occur only via a transition between a state where both patches are colonized and the extinction state, such that the MTE is given by r_{14}^{-1} . Here, the action can be found using the fact that the total metapopulation size is slowly varying compared to the local population size on each patch [19].

Dividing Hamiltonian (13) by μ we define

$$\tilde{H}(\mathbf{x}, \mathbf{p}) \equiv H_0(\mathbf{x}, \mathbf{p}) + \mu^{-1}H_1(\mathbf{x}, \mathbf{p}), \quad (26)$$

with

$$H_0(\mathbf{x}, \mathbf{p}) = x_1(e^{p_2 - p_1} - 1) + x_2\alpha(e^{p_1 - p_2} - 1),$$

$$H_1(\mathbf{x}, \mathbf{p}) = \sum_{i=1}^2 \{(e^{p_i} - 1)[b_i(x_i) - e^{-p_i}d_i(x_i)]\}. \quad (27)$$

That is, the local dynamics act as a perturbation to H_0 , which includes the migration terms. Assuming that the total population size varies much slower than that of each individual patch, we use the following transformation:

$$Q = x_1 + x_2, \quad q = x_2, \quad P = \frac{1}{2}(p_1 + p_2), \quad p = p_2. \quad (28)$$

This transformation, which is canonical in the leading order in $\mu \gg 1$, is motivated by the requirement that both Q and P are slowly varying compared to q and p . We now substitute this transformation into Hamiltonian (26) and write the Hamilton equations, $\dot{p} = -\partial_q \tilde{H}(Q, q, P, p)$ and $\dot{q} = \partial_p \tilde{H}(Q, q, P, p)$. Putting $\dot{p} = \dot{q} = 0$, i.e., assuming the

fast variables instantaneously equilibrate to some (Q, P) -dependent functions [42], and solving the resulting algebraic equations for q and p perturbatively with respect to μ^{-1} , we find

$$q = \frac{Q}{1+\alpha} + \mu^{-1}q^{(1)}(Q, P), \quad p = P + \mu^{-1}p^{(1)}(Q, P), \quad (29)$$

where $q^{(1)}$ and $p^{(1)}$ are (known) functions of Q and P . Using Eqs. (28) and (29) in Hamiltonian (26) divided by μ^{-1} , and keeping terms up to $O(\mu^{-1})$, we note that the zeroth-order term in μ vanishes, and we arrive at an approximate Hamiltonian in the fast migration regime:

$$H_{\text{eff}} = H_{\text{eff}}^{(0)}(Q, P) + \mu^{-1}H_{\text{eff}}^{(1)}(Q, P) + O(\mu^{-2}). \quad (30)$$

Here $H_{\text{eff}}^{(0)}(Q, P) = (e^P - 1)[\tilde{b}(Q) - e^{-P}\tilde{d}(Q)]$, $\tilde{b}(x) = 2x^2/[\tilde{\kappa}(1 + 1/\alpha)(1 - \tilde{\delta}^2)]$, $\tilde{d}(x) = x + x^3/[\tilde{\kappa}^2(1 + 1/\alpha)^2(1 - \tilde{\delta}^2)]$, while $H_{\text{eff}}^{(1)}(Q, P)$ is a (known) function of Q and P , but too long to be explicitly presented. To find the action, we equate H_{eff} [Eq. (30)] to zero, which yields

$$P(Q) = P^{(0)}(Q) + \mu^{-1}P^{(1)}(Q), \quad P^{(0)}(Q) = \ln[\tilde{d}(Q)/\tilde{b}(Q)]. \quad (31)$$

As a result, using Eqs. (28) and (29), the action satisfies

$$\begin{aligned} S_{\text{fast}} &= \int_{\xi_+ + O(\mu^{-1})}^{\xi_- + O(\mu^{-1})} p_1 dx_1 + \int_{\xi_+ / \alpha + O(\mu^{-1})}^{\xi_- / \alpha + O(\mu^{-1})} p_2 dx_2 \\ &\simeq \int_{Q_+ + O(\mu^{-1})}^{Q_- + O(\mu^{-1})} P^{(0)} dQ + \mu^{-1} \int_{Q_+}^{Q_-} \left(P^{(1)} + \frac{1-\alpha}{1+\alpha} p^{(1)} \right) dQ. \end{aligned} \quad (32)$$

Here and in Eq. (31) $P^{(1)}(Q)$ is too long to be explicitly presented, $Q_{\pm} = (1 + 1/\alpha)\xi_{\pm}$ is the combined population size of the two patches for the colonized state (Q_+) and for the colonization threshold (Q_-), while ξ_{\pm} are given by Eqs. (9).

Equation (32) allows computing S_{fast} up to subleading order in μ^{-1} . Indeed, the approximation in the second line of Eq. (32) contains two integrals: in the first we integrate over the zeroth-order trajectory, $P^{(0)}$, and take into account possible corrections to the limits, while in the second, we integrate over the corrections to the trajectory, but neglect corrections to the limits as these contribute only $O(\mu^{-2})$ terms to the action. To find the second integrand, we note that $P^{(1)}$ is a known function of Q , while $p^{(1)}(Q, P)$ must be evaluated at $P^{(0)}$, given by Eq. (31), as higher-order corrections in P contribute only $O(\mu^{-2})$ terms to the action. In what follows, we compute the zeroth-order term of S_{fast} as well as the $O(\mu^{-1})$ correction (in particular limits where the result is amenable).

The leading-order contribution to S_{fast} reads

$$S_{\text{fast}}^{(0)} = \int_{Q_+}^{Q_-} P^{(0)} dQ = (1 + 1/\alpha)\tilde{\kappa}S_0(\tilde{\delta}), \quad (33)$$

where $S_0(\delta)$ [Eq. (21)] is the action of an isolated patch. This result provides an indication whether fast migration is beneficial over isolation for the entire metapopulation. That is, if $S_{\text{fast}}^{(0)} > \max\{S_0(\delta_1), \kappa S_0(\delta_2)\}$ [43], fast migration has a positive effect on the population's viability. In this case, in addition to μ_{crit} for which the MTE is minimized, there exists an *optimal* migration rate, μ_{opt} , which globally maximizes

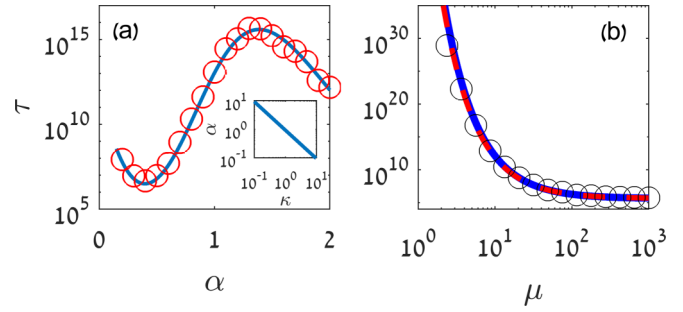


FIG. 5. (a) MTE in the fast migration limit as a function of α : theory [Eq. (32)] (solid line) versus WE simulations (circles). Inset shows the values of α and κ that maximize τ ; here $N = 300$, $\kappa = 2/3$, $\delta_1 = 0.56$, $\delta_2 = 0.34$, $\mu = 100$. (b) MTE in the fast migration limit as a function of μ : theory [Eq. (32)] (solid line), the $\kappa \ll 1$ approximation [Eqs. (34)] (dashed line), and WE simulations (circles). Here $N = 10^5$, $\kappa = 0.008$, $\delta_1 = 0.4$, $\delta_2 = 0.72$, and $\alpha = 1$.

the MTE; see Fig. 3(a) and below. In contrast, in Fig. 3(b) this condition does not hold, and τ decreases monotonically for the entire range of μ . Note that, upon replacing the threshold and carrying capacity parameters by their effective counterparts [see Eqs. (10)], $S_{\text{fast}}^{(0)}$ coincides with the one-patch result, up to a factor of $(1 + 1/\alpha)$, entailing the combined (deterministic) contribution of the two patches (see Sec. II).

Notably, $S_{\text{fast}}^{(0)}$ receives a maximum when $\alpha \simeq \kappa^{-1}$ (for comparable δ_1 and δ_2); see Fig. 5(a). That is, at $\mu \gg 1$, the extinction risk is minimized when the typical flux between patches is approximately equal, i.e., when typically an equal number of individuals pass across patches per unit time. This corresponds to an optimal synchronization of patches. In contrast, as $\alpha\kappa$ significantly deviates from 1, synchronization breaks down and one patch becomes much less stable than the other resulting in a dramatic decrease of the MTE. In extreme cases (see below) loss of synchrony may lead to “source-sink” dynamics where the patch with the large carrying capacity becomes a sink to the patch with the small carrying capacity.

1. Explicit calculation of correction

We now turn to discuss the correction term, $S_{\text{fast}}^{(1)}$, such that $S_{\text{fast}} = S_{\text{fast}}^{(0)} + \mu^{-1}S_{\text{fast}}^{(1)}$. While the full expression for $S_{\text{fast}}^{(1)}$ is not presented as it is highly cumbersome, our theoretical result in Figs. 3 and 5(b) includes this correction. In Fig. 6 we plot $S_{\text{fast}}^{(1)}$ as a function of δ_1 and δ_2 for two different parameter

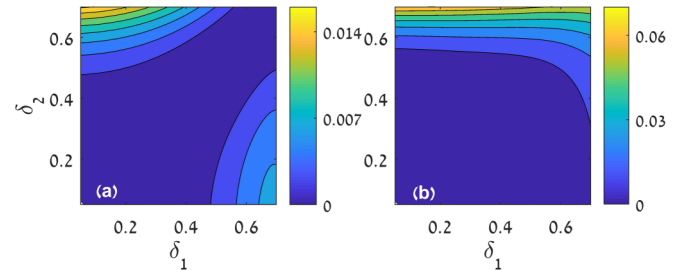


FIG. 6. $S_{\text{fast}}^{(1)}$ [the $O(\mu^{-1})$ correction in Eq. (32)] versus δ_1 and δ_2 . Here $\alpha = 2$ and $\kappa = \alpha^{-1} = 0.5$ (a) and $\kappa = 1$ (b). In (a), along the line $\delta_1 = \delta_2$, we find that $S_{\text{fast}}^{(1)} = 0$.

sets. We find that in most realistic cases $S_{\text{fast}}^{(1)}$ is positive, which suggests the existence of an optimal migration rate that maximizes the metapopulation's lifetime; see below.

In two limiting cases the correction $S_{\text{fast}}^{(1)}$ drastically simplifies. The first and simplest case is of well-mixed patches [19] $\kappa = 1/\alpha$ and $\delta_1 = \delta_2$. Here we find that the correction vanishes; see Fig. 6(a). This occurs since all terms that depend on μ^{-1} in Eq. (32), $P^{(1)}$, $p^{(1)}$ and the corrections to the integral limits, vanish in this special case. Thus, for well-mixed patches, the MTE rapidly reaches a steady value in the fast migration regime.

The second limit in which $S_{\text{fast}}^{(1)}$ simplifies is the limit of very different carrying capacities, $\kappa \ll 1$; here loss of synchrony may lead to a source-sink dynamics, where the large patch becomes a sink to the neighboring small patch. In Sec. II we have shown that deterministically, for $\kappa \ll 1$, the smaller patch dictates the typical size of the population. Here the leading and subleading-order terms in Eq. (32) become

$$\begin{aligned} S_{\text{fast}}^{(0)} &\simeq (\alpha + 1)\kappa S_0(\Delta_2), \\ S_{\text{fast}}^{(1)} &\simeq \alpha\kappa[4\text{arctanh}(\Delta_2) - 3\Delta_2 - 9S_0(\Delta_2)/2], \end{aligned} \quad (34)$$

where $S_0(\delta)$ is given by Eq. (21) and $\Delta_2 = [\delta_2^2 - \alpha(1 - \delta_2^2)]^{1/2}$. Thus, for $\kappa \ll 1$, Eqs. (34) indicate that the MTE is independent on the patch with the higher carrying capacity, i.e., the smaller patch dictates not only the metapopulation's typical size, but also its survival probability.

In Fig. 5(b) we compare for $\kappa \ll 1$ the general MTE [Eq. (32)] and Eqs. (34) and find very good agreement.

2. Optimal migration rate

As discussed above, in general, the optimal migration rate μ_{opt} , which maximizes the global MTE, exists if $S_{\text{fast}}^{(0)} > \max\{S_0(\delta_1), \kappa S_0(\delta_2)\}$, and if the correction $S_{\text{fast}}^{(1)}$ is positive. μ_{opt} can be found by a simple calculation, comparing the solutions for slow and fast migration:

$$\tau_{\text{slow}}(\mu_{\text{opt}}) = \tau_{\text{fast}}(\mu_{\text{opt}}), \quad (35)$$

where $\tau_{\text{slow}}(\mu_{\text{opt}})$ is the slow-migration MTE [Eq. (15)], and $\tau_{\text{fast}}(\mu_{\text{opt}}) \simeq \exp[NS_{\text{fast}}]$ can be found from Eq. (32).

In general, Eq. (35) can only be solved numerically. Yet, close to the bifurcation limit where $\delta_1, \delta_2 \ll 1$, Eqs. (15) and (32) drastically simplify, and we find $\mu_{\text{opt}} \sim 2\delta/3 + O[\delta(u-1)]$, where we have put $\delta_1 = \delta$, $\delta_2 = \delta u$ and we have assumed $|u-1| \ll 1$. The general solution of Eq. (35) and its comparison with the approximate solution close to bifurcation are shown in Figs. 7(a) and 7(b), respectively as function of κ and δ . In both panels we made sure that the correction $S_{\text{fast}}^{(1)}$ is positive such that there exists a maximum, apart from the special case of $\kappa = 1$ in Fig. 7(a) where $S_{\text{fast}}^{(1)} = 0$.

IV. WEIGHTED ENSEMBLE SIMULATIONS

Here we briefly discuss the weighted ensemble (WE) algorithm we have used to verify our analytic results and theoretical assumptions. Throughout this work, WE simulations were used to probe multiple transitions in space, and their associated probabilities. For example, in Fig. 8 we give snapshots from a WE simulation at different times. Here one observes the flow of probability from FP1 to both FP2 and

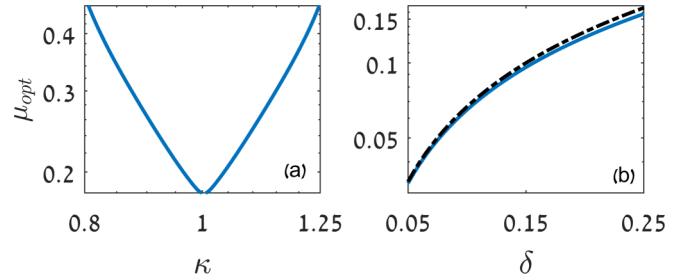


FIG. 7. (a) μ_{opt} given by numerically solving Eq. (35), for $\delta_1 = \delta_2 = 0.3$, $\alpha = 1$. (b) μ_{opt} given by numerically solving Eq. (35) (solid line) along with the result close to bifurcation (dashed-dotted line); see text. Here $\delta_1 = \delta_2 = \delta$, $\kappa = \alpha^{-1} = 7/11$.

FP3, and ultimately from FP2 and FP3 to FP4 [see Fig. 2(a)], where at long times the system reaches a quasi-stationary distribution of population sizes. As assumed in Sec. III A, this figure displays that extinction occurs serially, and that transitions like FP1 \rightarrow FP4 occur with very low probabilities.

The basic idea of the WE algorithm is to run significantly more simulations in regions of interest, whereas to compensate for the bias, we distribute the weight of each trajectory accordingly. To this end, space is divided into bins, which can be predefined or interactively chosen (on the fly), to ensure sampling in specific regions of interest. We thus start the simulation with m trajectories in proximity to a stable fixed point of the deterministic system. Each of the m trajectories is given an initial equal weight of $1/m$. The simulation consists of two general steps: (1) Trajectories are advanced in time by τ_{WE} , where the time-propagation method is according to the Gillespie algorithm [44]; (2) Trajectories are resampled as to maintain m trajectories in each occupied bin, while bins that are unoccupied remain so. An illustration of the method is given in Fig. 2(b), in which the number of bins is four and the number of trajectories is $m = 3$. The process of resampling can be done in various ways, as long as the distribution is maintained. In our simulation we used the original resampling method suggested by Huber and Kim [37].

Note that τ_{WE} is chosen to be much shorter than the system's relaxation time, $t_{i,+}^{\text{relax}}$, but much longer than the typical time between reactions, as to increase efficiency. We also stress that bins need to be chosen wisely: if chosen too

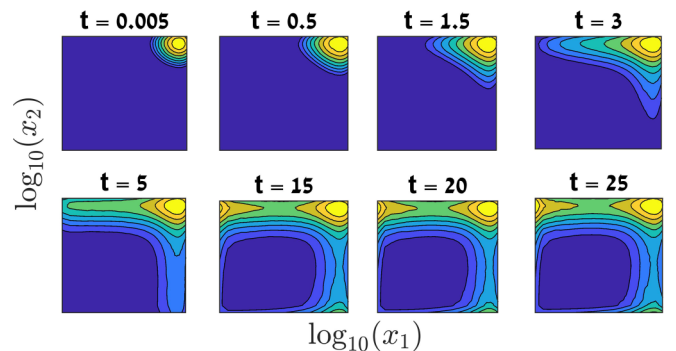


FIG. 8. Contour plot of the probabilities \mathbb{P}_{n_1, n_2} (see Sec. III) obtained by WE simulations. Parameters are $N = 100$, $\kappa = 1$, $\delta_1 = 0.5$, $\delta_2 = 0.4$, $\mu = 10^{-3}$, and $\alpha = 1$.

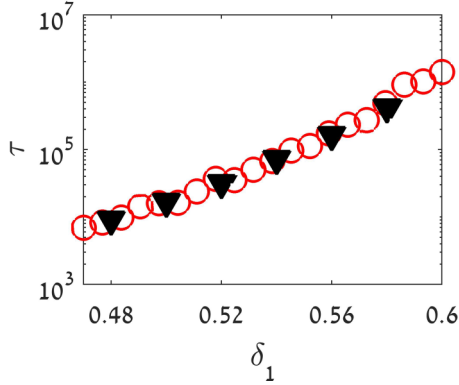


FIG. 9. MTE obtained by brute-force Monte Carlo (full triangles) and WE (circles) simulations, as a function of δ_1 . Parameters are $N = 60$, $\kappa = 1$, $\delta_2 = 0.45$, $\mu = 2$, and $\alpha = 1$.

far apart, trajectories will not reach remote regions, while if chosen close together the computational cost will be very high. Generally, there is a tradeoff between the number of bins and the trajectories per bin, assuming some memory limit. In our simulations, to achieve high efficiency we interactively changed the bins.

Error evaluation in our simulations was conducted numerically: by altering various simulation parameters, such as the number of bins and trajectories per bin, our estimation of the maximal error is $\sim 20\%$. This error is accounted for via the symbols' size in all relevant figures.

Finally, we checked that the results of the WE simulations coincide with brute-force Monte Carlo (BFMC) simulations in parameter regimes in which the latter are applicable; see e.g., Fig. 9. We stress that WE simulations are much more efficient than BFMC simulations, which are limited in probing rare events, as simulation times grow exponentially with the system's size. In addition, BFMC lacks the ability to easily separate different routes of extinction, which is at the center of this study. WE simulations are thus ideal for our purpose: larger system sizes do not require exponentially longer simulations and additionally, by measuring the flux between different metastable states, we can easily differentiate between global extinction and individual routes to extinction.

V. SUMMARY AND REALISTIC MODEL

In this work we have studied the extinction risk of a metapopulation—a network of interconnected population patches—where local patch dynamics include the Allee effect. We have shown that for slow migration the Allee effect gives rise to multiple routes to extinction, and generally decreases global survival probability. This may lead to the existence of a critical migration rate, which maximizes the global extinction risk of the population. These results are in sharp contrast to the case of local logistic dynamics [20]. Importantly, our results challenge the so-called rescue effect, where we have shown that as the migration rate increases, the global extinction risk also increases, even though each patch can separately experience a decrease in its local extinction risk.

In the fast migration regime, we have shown that there is only one route to extinction and found the MTE up to

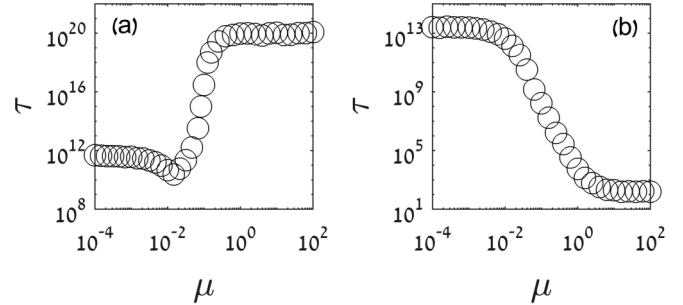


FIG. 10. MTE as a function of $\mu_{12} = \mu_{21} = \mu$: results of WE simulations for the alternative model exhibiting the Allee effect [Eq. (36)]. (a) $N_1 = 490$, $N_2 = 480$, and $n_{0,1} = n_{0,2} = 210$; (b) $N_1 = 500$, $N_2 = 70$, $n_{0,1} = 211$, and $n_{0,2} = 70$.

subleading-order corrections. Here, when the typical flux between patches is comparable, we have shown that migration is beneficial, and there usually exists an optimal migration rate that minimizes the extinction risk. Yet, in some cases isolation is preferred over mixing regardless of the migration rate, in contrast to local logistic dynamics [20]. This behavior occurs since under the Allee effect, patches do not always well mix [19].

All our analytical results were verified using a highly efficient WE method, which in general is very useful for studying rare events in stochastic population dynamics.

Notably, the analysis above is not limited to our particular choice of the birth and death rates [Eq. (1)], and is generic for any model, locally exhibiting the Allee effect [45]. To give another example we now present an alternative model which is more realistic in the biological sense; see, e.g., Refs. [19,45]. Here the Allee effect is locally accounted for by choosing the following birth-death process:

$$\begin{aligned} n_i &\xrightarrow{B_{n_i}} n_i + 1, & n_i &\xrightarrow{D_{n_i}} n_i - 1, \\ B_{n_i} &= \frac{N_i n_i^2}{n_i^2 + n_{0,i}^2}, & D_{n_i} &= n_i, \end{aligned} \quad (36)$$

where $N_i > 0$ is the carrying capacity, $0 < n_{0,i} < N_i/2$ is the threshold parameter, and migration between patch i and j occurs at a rate μ_{ij} . We have simulated this model for two patches using the WE simulations; see Sec. IV. Our results (see Fig. 10) indicate that the analysis done for the simple model $2A \leftrightarrow 3A$ and $A \rightarrow 0$ is generic and holds for other models exhibiting the Allee effect. In particular, our simulations demonstrate the existence of μ_{crit} and μ_{opt} in some parameter regimes [Fig. 10(a)], while in other regimes we observe a monotone decreasing MTE as a function of μ [Fig. 10(b)], similarly as in Fig. 3.

Notably, our approach may also be useful in analyzing a multistate gene regulatory network, where each “patch” corresponds to a distinct DNA state. Here the local Allee-like dynamics of proteins is supplemented by protein influx such that instead of extinction, the system switches between different phenotypic states [46–48]. Our method allows rigorous treatment of such models in the important limits of fast and slow binding or unbinding of a repressor or promoter to the DNA states, compared to protein dynamics [49]. Moreover,

our approach may provide insight into bet-hedging strategies of a bacterial population under antibiotic stress, where it has been observed that demographic fluctuations can be reduced by migration between the two “patches,” corresponding to the persister and nonpersister phenotypic states [50].

ACKNOWLEDGMENTS

We thank Michael Khasin and Yonatan Friedman for useful discussions. We acknowledge support from the Israel Science Foundation Grant No. 300/14 and the United States–Israel Binational Science Foundation Grant No. 2016-655. O.V. also acknowledges support from the Moscona Foundation.

APPENDIX A

In this Appendix we derive the MTE given in Eq. (15). Our starting point is Eqs. (14), which explicitly read

$$\begin{aligned}\dot{\mathcal{P}}_1(t) &= r_{21}\mathcal{P}_2(t) + r_{31}\mathcal{P}_3(t) - (r_{12} + r_{13} + r_{14})\mathcal{P}_1(t), \\ \dot{\mathcal{P}}_2(t) &= r_{12}\mathcal{P}_1(t) + r_{32}\mathcal{P}_3(t) - (r_{24} + r_{21} + r_{23})\mathcal{P}_2(t), \\ \dot{\mathcal{P}}_3(t) &= r_{13}\mathcal{P}_1(t) + r_{23}\mathcal{P}_2(t) - (r_{34} + r_{31} + r_{32})\mathcal{P}_3(t), \\ \dot{\mathcal{P}}_4(t) &= r_{14}\mathcal{P}_1(t) + r_{24}\mathcal{P}_2(t) + r_{34}\mathcal{P}_3(t),\end{aligned}\quad (\text{A1})$$

where \mathcal{P}_i is the probability to be at the basin of attraction of FP i (see Fig. 2) and r_{ij} is the transition rate between \mathcal{P}_i and \mathcal{P}_j . Using the last of Eqs. (A1), the MTE, given by $\tau = \int_0^\infty t\mathcal{P}_4 dt$, reads

$$\tau = \int_0^\infty t[r_{14}\mathcal{P}_1(t) + r_{24}\mathcal{P}_2(t) + r_{34}\mathcal{P}_3(t)]dt. \quad (\text{A2})$$

Now, since $\mathcal{P}_4(t)$ does not appear explicitly in Eqs. (A1), we define $\mathcal{P}(t) = [\mathcal{P}_1(t), \mathcal{P}_2(t), \mathcal{P}_3(t)]$ and rewrite the first three of Eqs. (A1) in matrix form:

$$\dot{\mathcal{P}}(t) = \overleftrightarrow{A}\mathcal{P}(t), \quad (\text{A3})$$

with

$$\overleftrightarrow{A} \equiv \begin{pmatrix} -\sum_{i \neq 1}^4 r_{1i} & r_{21} & r_{31} \\ r_{12} & -\sum_{i \neq 2}^4 r_{2i} & 0 \\ r_{13} & 0 & -\sum_{i \neq 3}^4 r_{3i} \end{pmatrix}. \quad (\text{A4})$$

Note that in matrix \overleftrightarrow{A} we have neglected the rates r_{14} , r_{23} , and r_{32} . Since at slow migration extinction occurs in a serial manner, it can be shown that in the semiclassical limit where the carrying capacities are large, these rates are negligible compared to other rates, as they satisfy $r_{14} \sim r_{12}r_{24}$, $r_{23} \sim r_{21}r_{13}$, and $r_{32} \sim r_{31}r_{12}$; see below.

In order to find the MTE we need to solve (A3) with initial conditions $\mathcal{P}(0) = (1, 0, 0)$, i.e., starting from the colonized state in both patches. Since \overleftrightarrow{A} is not necessarily diagonalizable, the problem can be generally solved via the Schur decomposition [51], namely, $\overleftrightarrow{A} = \overleftrightarrow{U}\overleftrightarrow{T}\overleftrightarrow{U}^{-1}$, with \overleftrightarrow{U} being a unitary matrix and \overleftrightarrow{T} an upper triangular matrix, given by

$$\overleftrightarrow{T} \equiv \begin{pmatrix} \lambda_1 & T_{12} & T_{13} \\ 0 & \lambda_2 & T_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix}. \quad (\text{A5})$$

Here the diagonal elements $\lambda_i < 0$, are the eigenvalues of \overleftrightarrow{A} , and T_{12}, T_{13}, T_{23} depend on the chosen (not unique) Schur decomposition. By decomposing $\mathcal{P}(t) = \overleftrightarrow{U}\overleftrightarrow{q}(t)$ we rewrite Eq. (A3) as $\dot{\overleftrightarrow{q}}(t) = \overleftrightarrow{T}\overleftrightarrow{q}(t)$, with initial conditions $\overleftrightarrow{q}(0) = \overleftrightarrow{U}^{-1}\mathcal{P}(0)$. Since \overleftrightarrow{T} is upper triangular, the solution to $\overleftrightarrow{q}(t) = [q_1(t), q_2(t), q_3(t)]$ can be found iteratively by solving for $q_3(t)$, then $q_2(t)$, and then $q_1(t)$, as follows:

$$\begin{aligned}q_3(t) &= q_3(0)e^{\lambda_3 t}, \\ q_2(t) &= q_2(0)e^{\lambda_2 t} + \int_0^t T_{23}e^{-\lambda_2(t-t')}q_3(t')dt', \\ q_1(t) &= q_1(0)e^{\lambda_1 t} + \int_0^t e^{-\lambda_1(t-t')}[T_{12}q_2(t') + T_{13}q_3(t')]dt',\end{aligned}\quad (\text{A6})$$

where these integrals can be computed in a straightforward manner. Finally, having found $\overleftrightarrow{q}(t)$, Eq. (A2) is solved by

$$\tau = \mathbf{r}_4^T \overleftrightarrow{U} \mathbf{c}, \quad (\text{A7})$$

where $\mathbf{r}_4 = (0, r_{24}, r_{34})$ and $\mathbf{c} = \int_0^\infty t\overleftrightarrow{q}(t)dt$ ($i = 1, 2, 3$). Note that while the Schur decomposition and the solution for \overleftrightarrow{q} are not unique, the solutions for \mathcal{P} and τ [Eq. (A7)] are in fact unique.

While an explicit version of Eq. (A7) exists, it is highly cumbersome, as it involves solving a cubic equation and finding a specific schur decomposition. Nonetheless, in one important case τ drastically simplifies. In general, if the carrying capacities are large, the transition rates exponentially differ from each other. This is due to the large parameter N that multiplies the action in the exponent, and the fact that the actions are all, in general, different from one another. In particular, this indicates that one of the transition rates $\text{FP1} \rightarrow \text{FP2}$ or $\text{FP1} \rightarrow \text{FP3}$, respectively given by r_{12} and r_{13} (see Fig. 2), is negligible compared to the other. Without loss of generality, let us assume that r_{13} is negligible compared to r_{12} . In this case, the solution to (A7) drastically simplifies and reads

$$\tau = \frac{r_{12} + r_{21} + r_{24}}{r_{12}r_{24}} \simeq \max\left\{\frac{1}{r_{24}}, \frac{1}{r_{12}}, \frac{r_{21}}{r_{12}r_{24}}\right\}. \quad (\text{A8})$$

While this result accounts for only one route to extinction (assuming the other route is exponentially unlikely), it also includes a correction to this single route due to possible recolonization, which depends on the colonization rate r_{21} . Demanding *a posteriori* that r_{13} be negligible compared to the individual rates comprising this route to extinction, and using Eq. (20), it can be shown that Eq. (A8) is valid as long as $r_{13} \ll r_{12}r_{24}/r_{21}$; otherwise we have neglected a term larger than the rate of extinction. Note that if alternatively $r_{13} \gg r_{12}$ then the equivalent of Eq. (A8) will hold as long as $r_{12} \ll r_{13}r_{34}/r_{31}$. Also note that these conditions are easily satisfied when the colonization rates r_{21} and r_{31} are negligibly small, which is the case in most of the parameter space, whereas Eq. (A8) breaks down only in non-WKB parameter regimes. These arguments lead to a generic expression for τ , given by Eq. (15).

We finally show, by explicitly calculating r_{14} , that for slow migration extinction occurs in a serial manner with an overwhelming probability. In the leading order in μ , the action is the sum of all independent actions [20]:

$$S_{14} = \sum_{i=1}^2 \int_{(x_{+,+})_i}^{(x_{-,-})_i} p_i dx_i \simeq S_0(\delta_1) + \kappa S_0(\delta_2). \quad (\text{A9})$$

Since in the limit $\mu \rightarrow 0$, this expression coincides with the sum of the actions S_{12} and S_{24} or S_{13} and S_{34} , we find that $r_{14} \sim r_{12}r_{24} \sim r_{13}r_{34}$, which means that r_{14} is exponentially smaller than any of the individual rates. Clearly, the same argument also applies for rates r_{23} and r_{32} which include transitions resulting in changes in both patches. Moreover, the fact that r_{14} , r_{23} , and r_{32} , are negligibly small compared to all other rates, was also verified by WE simulations. Note, that similarly as done in Ref. [20], subleading-order corrections in $\mu \ll 1$ can also be computed for r_{14} . Yet we do not give the corrections here as r_{14} itself is negligible in the slow migration regime.

APPENDIX B

Here we generalize our results to M fully connected patches, for both slow and fast migration.

1. The case of slow migration

The results for two patches [Eqs. (20)] can be generalized to M connected patches. This is done by defining $\phi_i^{\text{in}} = \sum_{j \neq i} \alpha_j x_{j,s_j}^{(0)}$ [see Eq. (5)] as the incoming flux to patch i given by the sum over all zeroth-order stable FPs of patches migrating into this patch [$s_j = (0, +)$]. Furthermore, defining $\phi_i^{\text{out}} = \sum_{j \neq i} \alpha_j$ as the magnitude of the outgoing flux from patch i , and using similar arguments as in the two-patch case, we find

$$S_i(\phi_i^{\text{in}}) = \kappa_i S_0(\delta_i)(1 - \mu \phi_i^{\text{out}}/2) + \mu \delta_i (\phi_i^{\text{in}} - \phi_i^{\text{out}} \kappa_i), \quad (\text{B1})$$

where $S_0(\delta_i)$ is given by Eq. (21). Since ϕ_i^{in} depends on $x_{j,s_j}^{(0)}$ for $j \neq i$ and $s_j = (0, +)$, ϕ_i^{in} can take 2^{M-1} different values, as each of the other patches can be either extinct or colonized. On the other hand, ϕ_i^{out} is determined only by patch i . The notation $S_i(\phi_i^{\text{in}})$ is thus intended to emphasize that this action describes the extinction of patch i , given a specific influx ϕ_i^{in} , chosen out of 2^{M-1} possibilities for the deterministic occupancy of the other patches.

Similarly, the colonization rate of patch i , given by Eqs. (22) and (23), can also be generalized to M patches:

$$\begin{aligned} S_i(\phi_i^{\text{in}}) = & \kappa_i \left\{ \delta_i + 2(1 - \delta_i^2)^{1/2} \arcsin \left[\left(\frac{1 - \delta_i}{2} \right)^{1/2} \right] - 1 \right\} \\ & - \mu^{1/2} \frac{\pi [\phi_i^{\text{in}} (1 - \delta_i^2) \kappa_i]^{1/2}}{2^{1/2}} \\ & + \mu \left\{ \phi_i^{\text{out}} (1 - \delta_i^2)^{1/2} \kappa_i \arcsin \left[\left(\frac{1 - \delta_i}{2} \right)^{1/2} \right] \right. \\ & \left. + \frac{1}{2} \phi_i^{\text{in}} (\delta_i + 3) \right\}, \end{aligned} \quad (\text{B2})$$

where in the case of two patches, S_{31} and S_{21} are obtained from this result for $i = 1$, $\phi_1^{\text{out}} = 1$, $\phi_1^{\text{in}} = \alpha \kappa (1 + \delta_2)$ and $i = 2$, $\phi_2^{\text{out}} = \alpha$, $\phi_2^{\text{in}} = (1 + \delta_1)$, respectively. Using similar arguments to those in the case of two patches, one can show that our main result—that for sufficiently slow migration, mixing reduces stability—also holds for the general case of M patches. To test this claim, we have performed additional numerical WE simulations (not shown) with up to four interconnected patches.

2. The case of fast migration

The zeroth-order result in the fast migration limit, Eq. (33), can also be generalized to M patches. Here we choose for simplicity $\mu_{ij} = \mu$. By conducting a similar calculation to the two-patch case one finds $S_{\text{fast}}^{(0)} = M \tilde{\kappa} S_0(\tilde{\delta})$, where here $\tilde{\kappa}$ and $\tilde{\delta}$ are given by [19]

$$\begin{aligned} \frac{M}{\tilde{\kappa}(1 - \tilde{\delta}^2)} &= \sum_{i=1}^M \frac{1}{\kappa_i(1 - \delta_i^2)}, \\ \frac{M}{\tilde{\kappa}^2(1 - \tilde{\delta}^2)} &= \sum_{i=1}^M \frac{1}{\kappa_i^2(1 - \delta_i^2)}, \end{aligned} \quad (\text{B3})$$

and it is required that $\tilde{\delta}$ be real, as in the two-patch case. Notably, by numerically analyzing this result one can show that synchronization maximizes stability also for the general case of M interconnected patches. Namely, for $\mu_{ij} = \mu$, the action receives a maximum when the carrying capacities of all patches are approximately equal. We have verified this result by performing numerical WE simulations (not shown) with up to four interconnected patches.

- [1] I. A. Hanski and O. E. Gaggiotti, *Ecology, Genetics and Evolution of Metapopulations* (Academic Press, Cambridge, 2004).
 [2] L. Fahrig, *Annu. Rev. Ecol. Evol. System.* **48**, 1 (2017).
 [3] B. A. Wilcox and D. D. Murphy, *Am. Nat.* **125**, 879 (1985).
 [4] R. Levins, *Am. Entomol.* **15**, 237 (1969).
 [5] V. Colizza and A. Vespignani, *Phys. Rev. Lett.* **99**, 148701 (2007).
 [6] Y. M. Lai, J. Newby, and P. C. Bressloff, *Phys. Rev. Lett.* **107**, 118102 (2011).

- [7] S. Fedotov and H. Stage, *Phys. Rev. Lett.* **118**, 098301 (2017).
 [8] P. A. Stephens, W. J. Sutherland, and R. P. Freckleton, *Oikos* **87**, 185 (1999).
 [9] R. Lande, *Am. Nat.* **130**, 624 (1987).
 [10] P. Amarasekare, *Am. Nat.* **152**, 298 (1998).
 [11] R. Lande, S. Engen, and B.-E. Sæther, *Oikos* **83**, 383 (1998).
 [12] I. Hanski and O. Ovaskainen, *Nature (London)* **404**, 755 (2000).
 [13] O. Ovaskainen and I. Hanski, *Theor. Pop. Biol.* **60**, 281 (2001).
 [14] M. Kuussaari, I. Saccheri, M. Camara, and I. Hanski, *Oikos* **82**, 384 (1998).

- [15] A. M. Kramer, B. Dennis, A. M. Liebhold, and J. M. Drake, *Pop. Ecol.* **51**, 341 (2009).
- [16] J. Sun, S. P. Cornelius, J. Janssen, K. A. Gray, and A. E. Motter, *J. R. Soc., Interface* **12**, 20150235 (2015).
- [17] B. Dennis, *Oikos* **96**, 389 (2002).
- [18] C. M. Taylor and A. Hastings, *Ecol. Lett.* **8**, 895 (2005).
- [19] M. Khasin, E. Khain, and L. M. Sander, *Phys. Rev. Lett.* **109**, 248102 (2012).
- [20] M. Khasin, B. Meerson, E. Khain, and L. M. Sander, *Phys. Rev. Lett.* **109**, 138104 (2012).
- [21] A. Eriksson, F. Elías-Wolff, B. Mehlig, and A. Manica, *Proc. R. Soc. B* **281**, 20133127 (2014).
- [22] M. Assaf and B. Meerson, *J. Phys. A: Math. Theor.* **50**, 263001 (2017).
- [23] O. Ovaskainen, *Ann. Zool. Fenn.* **54**, 113 (2017).
- [24] Here $(x_s)_i$ refers to element $i = 1, 2$ of vector x_s .
- [25] Similarly, for $\kappa \gg 1$, the metapopulation size is dictated by patch 1, and bistability requires $\delta_1 \geq [1/(1 + \alpha)]^{1/2}$.
- [26] M. Dykman, E. Mori, J. Ross, and P. Hunt, *J. Chem. Phys.* **100**, 5735 (1994).
- [27] M. Assaf and B. Meerson, *Phys. Rev. Lett.* **97**, 200602 (2006).
- [28] D. A. Kessler and N. M. Shnerb, *J. Stat. Phys.* **127**, 861 (2007).
- [29] B. Meerson and P. V. Sasorov, *Phys. Rev. E* **78**, 060103(R) (2008).
- [30] C. Escudero and A. Kamenev, *Phys. Rev. E* **79**, 041149 (2009).
- [31] M. Assaf and B. Meerson, *Phys. Rev. E* **81**, 021116 (2010).
- [32] M. I. Dykman, I. B. Schwartz, and A. S. Landsman, *Phys. Rev. Lett.* **101**, 078101 (2008).
- [33] I. B. Schwartz, L. Billings, M. Dykman, and A. Landsman, *J. Stat. Mech.: Theor. Exp.* (2009) P01005.
- [34] O. Gottesman and B. Meerson, *Phys. Rev. E* **85**, 021140 (2012).
- [35] Here we assume that $r_{ai} = 0$ for $i = 1, 2, 3$ since the global extinction state (FP4) is absorbing.
- [36] Note that Eqs. (22) and (23) contain a $O(\mu^{1/2})$ term. This occurs since when integrating over the optimal path [Eq. (17)], the lower boundary scales with μ . It can thus be checked that the contribution from the integral over the optimal path, up to $x = O(\mu^{1/2})$, is not negligible.
- [37] G. A. Huber and S. Kim, *Biophys. J.* **70**, 97 (1996).
- [38] B. W. Zhang, D. Jasnow, and D. M. Zuckerman, *J. Chem. Phys.* **132**, 054107 (2010).
- [39] D. M. Zuckerman and L. T. Chong, *Annu. Rev. Biophys.* **46**, 43 (2017).
- [40] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, *Nature (London)* **461**, 53 (2009).
- [41] We have numerically confirmed that for a reasonable choice of parameters, for δ_i not too close to 1, the colonization rate of patch i is negligible compared to its extinction rate (given that the other patch is colonized). This certainly holds close to bifurcation, where $\delta_i \ll 1$.
- [42] M. Assaf and B. Meerson, *Phys. Rev. Lett.* **100**, 058105 (2008).
- [43] In the case where the patches are isolated, the MTE is determined by the higher action of the individual patches.
- [44] D. T. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).
- [45] V. Méndez, M. Assaf, A. Masó-Puigdellosas, D. Campos, and W. Horsthemke, *Phys. Rev. E* **99**, 022101 (2019).
- [46] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie, *Science* **322**, 442 (2008).
- [47] M. Assaf, E. Roberts, and Z. Luthey-Schulten, *Phys. Rev. Lett.* **106**, 248102 (2011).
- [48] T. M. Earnest, E. Roberts, M. Assaf, K. Dahmen, and Z. Luthey-Schulten, *Phys. Biol.* **10**, 026002 (2013).
- [49] M. J. Morelli, P. R. ten Wolde, and R. J. Allen, *Proc. Natl. Acad. Sci. USA* **106**, 8101 (2009).
- [50] S. Pearl, C. Gabay, R. Kishony, A. Oppenheim, and N. Q. Balaban, *PLoS Biol.* **6**, e120 (2008).
- [51] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 2012).