The Role of Stochasticity in Noise-Induced Tipping Point Cascades: A Master Equation Approach

Abhishek Mallela1 · Alan Hastings2,3

Received: 18 November 2020 / Accepted: 12 March 2021 / Published online: 31 March 2021
© The Author(s), under exclusive licence to Society for Mathematical Biology 2021

Abstract
Tipping points have been shown to be ubiquitous, both in models and empirically in a range of physical and biological systems. The question of how tipping points cascade through systems has been less explored and is an important one. A study of noise-induced tipping, in particular, could provide key insights into tipping cascades. Here, we consider a specific example of a simple model system that could have cascading tipping points. This model consists of two interacting populations with underlying Allee effects and stochastic dynamics, in separate patches connected by dispersal, which can generate bistability. From an ecological standpoint, we look for rescue effects whereby one population can prevent the collapse of a second population. As a way to investigate the stochastic dynamics, we use an individual-based modeling approach rooted in chemical reaction network theory. Then, using continuous-time Markov chains and the theory of first passage times, we essentially approximate, or emulate, the original high-dimensional model by a Markov chain with just three states, where each state corresponds to a combination of population thresholds. Analysis of this reduced model shows when the system is likely to recover, as well as when tipping cascades through the whole system.

Keywords Tipping points · Stochasticity · Allee effects · Alternative stable states · Resilience · First passage time

1 Introduction
Many systems in nature can transition into a qualitatively distinct dynamic regime when a critical threshold is approached. The associated threshold of such a system,
defined in the context of bifurcation theory, is the bifurcation point or tipping point of the system. Tipping points manifest in systems including lake eutrophication (ecology), social contagion (sociology), disease spread (epidemiology), epileptic seizures (physiology), stock market crashes (finance), and even the earth system (Scheffer 2009; Scheffer et al. 2012; O’Regan 2018; Lenton 2020; Klose et al. 2020). Tipping points arise in the presence of strongly self-amplifying (mathematically positive) feedbacks (Lenton 2020). This characterization of tipping points is important in the sense that only sufficiently strong, self-propelling feedbacks are recognized.

A prime example of a tipping point is an ecological system with an Allee effect (Courchamp et al. 2008; Hastings and Gross 2012; Johnson and Hastings 2018; Vortkamp et al. 2020). Allee effects occur in populations with low abundances and are believed to be commonplace in ecological systems (Courchamp et al. 2008; Drake and Kramer 2011). In a population exhibiting an Allee effect, the per capita growth rate is a unimodal function of the population abundance with a global maximum. The sign of the growth rate at low population levels distinguishes a weak Allee effect from a strong one. In particular, a weak Allee effect does not result in a negative growth rate for small population sizes but its strong counterpart presents with a negative rate. Therefore, the existence of a strong Allee effect implies the existence of a critical threshold for survival. The possibility of alternative stable states in systems that are analogous to those with a strong Allee effect has significant implications at many levels, from the microscopic scale of budding yeast (Dai et al. 2012) to the macroscopic scale of tipping elements for the Earth system (Lenton 2020; Klose et al. 2020).

In order to investigate how tipping points manifest in systems with a strong Allee effect, one needs to formulate a model that can exhibit this behavior. A fundamental question in this setting concerns the propagation of tipping points through an ecosystem consisting of multiple patches. To what extent are interacting populations interdependent and how is this relationship influenced by the parameters governing the model behavior? Stated more simply, how do tipping points cascade through systems?

A simple analysis considers multiple populations in a network that are connected by passive, symmetric diffusion. In this work, we study a model consisting of two populations to analyze how their dynamics are related, conditional on the quality of their internal environments. One case of interest is when one population has fallen below the Allee threshold and we ask whether the next transition puts both populations above, or alternatively, below, the Allee threshold. A similar study was conducted in a deterministic setting (Johnson and Hastings 2018). That study laid the foundations for an eventual treatment in a stochastic setting. Hence, in our work, we adopt the definition of ecological resilience formulated by Holling (1973) as well as the framework proposed by Johnson and Hastings, but we also account for stochasticity in the system.

The literature on single, isolated tipping points is vast. However, studies of cascades of tipping points are less common. We begin with the background needed to formulate a stochastic, individual-based model that accounts for strong Allee effects. Our model follows a Markovian birth–death process, inspired by the theory of chemical reaction networks. We find that this representation naturally lends itself to a treatment with Markov chains and first passage times. Then, we employ dimensionality reduction
techniques and novel approximations to propose a reduced model that we study in detail.

2 Model Description

Much of the work done on modeling deterministic systems with one population that manifest Allee effects are based on either empirical observations or phenomenology. The crux of these models dates back to 1954 (Odum and Allee 1954), where the observed per-capita growth rate was fit with a suitable function. The general form of such a deterministic model in continuous time is the following ordinary differential equation (ODE):

$$\frac{d\rho(t)}{dt} = \rho f(\rho),$$  \hspace{1cm} (1)

where \(\rho(t)\) denotes the average population density at time \(t\), and \(f(\rho)\) is a function specifying the form of the per-capita population growth rate at density \(\rho\). There have been many functional forms proposed for \(f(\rho)\) in the literature. A review of various specifications used in models can be found (Boukal and Berec 2002).

In particular, a simple and important model was proposed early in the twentieth century (Volterra 1938), where \(f(\rho)\) is a quadratic polynomial function of \(\rho\). The underlying assumptions of the model are as follows. Given a constant sex ratio, the number of meetings between males and females is proportional to \(\rho^2\). The ratio of births to meetings can be affected by the population density and is hence assumed to be linearly decreasing in \(\rho\). Also, the mortality rate of individuals in the population is assumed to be proportional to \(\rho\). Hence, the Volterra model takes the following form:

$$\frac{d\rho(t)}{dt} = -a_1 \rho + (a_2 - a_3 \rho) \rho^2 = -a_1 \rho + a_2 \rho^2 - a_3 \rho^3,$$  \hspace{1cm} (2)

where \(a_1, a_2, a_3 > 0\). If we define the two real-valued roots,

$$k_1 = \frac{1}{2a_3} \left[ a_2 - \sqrt{a_2^2 - 4a_1a_3} \right],$$  \hspace{1cm} (3a)

$$k_2 = \frac{1}{2a_3} \left[ a_2 + \sqrt{a_2^2 - 4a_1a_3} \right],$$  \hspace{1cm} (3b)

with \(a_2^2 > 4a_1a_3\), then the model is often shown in the following form (Courchamp et al. 1999)

$$\frac{d\rho(t)}{dt} = a_1 \rho \left( 1 - \frac{\rho}{k_2} \right) \left( \frac{\rho}{k_1} - 1 \right),$$  \hspace{1cm} (4)

resembling the logistic model with the addition of a new unstable steady state, \(\rho = k_1\).
For the strong Allee effect, the Volterra model has three steady states: two stable states at $\rho = 0$ and $\rho = k_2$ and an unstable state at $\rho = k_1$. If the initial population size exceeds $k_1$, the population grows over time and converges to the stable steady state $\rho = k_2$, the carrying capacity of the system. If the initial population size is less than $k_1$, the population decays over time to the stable steady state $\rho = 0$ and goes extinct.

In order to investigate the effects of internal fluctuations or demographic stochasticity in a two-population system with both weak and strong Allee effects, we consider the temporal evolution of the system as specified by a Markovian birth–death process. Demographic stochasticity is included both in the dynamics of the locations and in the dispersal parameter. We account for the individual reaction kinetics explicitly, in a mechanistic manner, without relying on phenomenological considerations. Here, we follow the approach of Méndez et al. (2019) by casting the system as a chemical reaction network that results in an individual-based model (IBM). The minimal IBM that displays both the weak and strong Allee effect and also accounts for dispersal can be described as follows. It consists of two birth processes (linear and binary birth), a ternary competition process, a linear death process, and an exchange process. We provide our reaction scheme below.

\begin{align}
X_1 & \xrightarrow{\mu_1} (1 + b)X_1 \quad (5a) \\
2X_1 & \xrightarrow{\lambda_1} (2 + a)X_1 \quad (5b) \\
X_1 & \xrightarrow{\gamma_1} \emptyset \quad (5c) \\
3X_1 & \xrightarrow{r_1} (3 - c)X_1 \quad (5d) \\
X_1 & \xrightarrow{d} X_2 \quad (5e) \\
X_2 & \xrightarrow{\mu_2} (1 + b)X_2 \quad (5f) \\
2X_2 & \xrightarrow{\lambda_2} (2 + a)X_2 \quad (5g) \\
X_2 & \xrightarrow{\gamma_2} \emptyset \quad (5h) \\
3X_2 & \xrightarrow{r_2} (3 - c)X_2 \quad (5i) \\
X_2 & \xrightarrow{d} X_1 \quad (5j)
\end{align}

The first reaction is a linear birth process, which occurs at a constant rate $\mu_1$, and describes the baseline reproductive success of the first population in the absence of cooperative effects. It accounts for the fact that the typical individual produces $b$ offspring that reach reproductive age. The second reaction is a binary process that occurs at a constant rate $\lambda_1$. It describes cooperative interactions, such as breeding, antipredator behavior, or foraging, that result in producing $a$ additional offspring which reach reproductive age. The third reaction is a linear death process, occurring at constant rate $\gamma_1$, which accounts for mortality due to natural causes. The fourth reaction is a ternary competition process, accounting for the results of overcrowding and resource deple-
tion, where \( c \) individuals die at rate \( \tau_1 \). Note that \( c = 1, 2, 3 \) are the only meaningful values. The next reaction is an exchange process of symmetric dispersal between the two populations. This occurs at a constant rate of \( d \). The last five reactions in the scheme describe the dynamics of the second population, respectively.

The reaction scheme (5) defines a Markovian process, and the temporal evolution of \( P(n_1, n_2, t) \), the probability of having \( n_i \) individuals from the \( i \)th population at time \( t \) for \( i = 1, 2 \), is described by the following master equation, also known as the forward Kolmogorov equation (Gardiner 2004):

\[
\frac{d P(n, t)}{dt} = \sum_r [W(n - r, r) P(n - r, t) - W(n, r) P(n, t)],
\]

where \( P(n < 0, t) = P(n_1 < 0, n_2 < 0, t) = 0 \). Here, \( W(n, r) \) denotes the set of transition rates between the states with \( n \) and \( n + r \) individuals, where

\[
r = \{r_1, r_2, \ldots, r_{10}\} = \{(b, 0), (a, 0), (-1, 0), (-c, 0), (-1, 0), (0, b), (0, a), (0, -1), (0, -c), (0, -1)\}
\]

is the vector of transition increments corresponding to the system given by Eq. (5). The transition rates corresponding to each reaction, \( W(n, r) \), are obtained from the reaction kinetics (Van Kampen 1992; Gardiner 2004):

\[
\begin{align*}
W(n, r_1) &= \mu_1 n_1, \\
W(n, r_2) &= \frac{\lambda_1}{2} n_1 (n_1 - 1), \\
W(n, r_3) &= \gamma_1 n_1, \\
W(n, r_4) &= \frac{\tau_1}{6} n_1 (n_1 - 1)(n_1 - 2), \\
W(n, r_5) &= dn_2, \\
W(n, r_6) &= \mu_2 n_2, \\
W(n, r_7) &= \frac{\lambda_2}{2} n_2 (n_2 - 1), \\
W(n, r_8) &= \gamma_2 n_2, \\
W(n, r_9) &= \frac{\tau_2}{6} n_2 (n_2 - 1)(n_2 - 2), \\
W(n, r_{10}) &= dn_1
\end{align*}
\]

Deterministic ODEs for the average population sizes can be obtained from Eq. (6). Multiplying Eq. (6) by \( n_1 n_2 \), using transition rates (7), and summing over all values of \( n_1 \) and \( n_2 \), we find
\[
\frac{d\rho_1}{dt} = (b\mu_1 - \gamma_1 - d)\rho_1 + \frac{a\lambda_1}{2} \rho_1^2 - \frac{c\tau_1}{6} \rho_1^3 + d\rho_2, \quad (8a)
\]
\[
\frac{d\rho_2}{dt} = (b\mu_2 - \gamma_2 - d)\rho_2 + \frac{a\lambda_2}{2} \rho_2^2 - \frac{c\tau_2}{6} \rho_2^3 + d\rho_1, \quad (8b)
\]

where \(\rho_i = \langle n_i \rangle\) is the mean number of individuals in population \(i\) for \(i = 1, 2\). We note that this pair of deterministic equations holds strictly when the demographic fluctuations vanish. This occurs in the macroscopic limit as the population size and spatial extent of each population increase to infinity such that their ratio, the population density, stays constant or approaches a finite limit.

For the sake of simplicity, in what follows, we focus on the simplest version of this IBM. Namely, we treat the Markovian process as a single-step process with \(a = b = c = 1\). The set of reactions (5) then becomes

\[
\begin{align*}
X_1 &\stackrel{\mu_1}{\rightarrow} 2X_1 \quad (9a) \\
2X_1 &\stackrel{\lambda_1}{\rightarrow} 3X_1 \quad (9b) \\
X_1 &\stackrel{\gamma_1}{\rightarrow} \emptyset \quad (9c) \\
3X_1 &\stackrel{\tau_1}{\rightarrow} 2X_1 \quad (9d) \\
X_1 &\stackrel{d}{\rightarrow} X_2 \quad (9e) \\
X_2 &\stackrel{\mu_2}{\rightarrow} 2X_2 \quad (9f) \\
2X_2 &\stackrel{\lambda_2}{\rightarrow} 3X_2 \quad (9g) \\
X_2 &\stackrel{\gamma_2}{\rightarrow} \emptyset \quad (9h) \\
3X_2 &\stackrel{\tau_2}{\rightarrow} 2X_2 \quad (9i) \\
X_2 &\stackrel{d}{\rightarrow} X_1 \quad (9j)
\end{align*}
\]

The mean-field rate equations corresponding to (9) are

\[
\begin{align*}
\frac{d\rho_1}{dt} &= (\mu_1 - \gamma_1 - d)\rho_1 + \frac{\lambda_1}{2} \rho_1^2 - \frac{\tau_1}{6} \rho_1^3 + d\rho_2, \quad (10a) \\
\frac{d\rho_2}{dt} &= (\mu_2 - \gamma_2 - d)\rho_2 + \frac{\lambda_2}{2} \rho_2^2 - \frac{\tau_2}{6} \rho_2^3 + d\rho_1 \quad (10b)
\end{align*}
\]

For this set of reactions, the master equation can be explicitly obtained as

\[
\begin{align*}
\frac{dP(n_1, n_2, t)}{dt} &= (n_1 - 1) \left[ \frac{\lambda_1}{2} (n_1 - 2) + \mu_1 \right] P(n_1 - 1, n_2, t) \\
&\quad + (n_1 + 1) \left[ \frac{\tau_1}{6} n_1 (n_1 - 1) + \gamma_1 + d \right] P(n_1 + 1, n_2, t) \\
&\quad + (n_2 - 1) \left[ \frac{\lambda_2}{2} (n_2 - 2) + \mu_2 \right] P(n_1, n_2 - 1, t)
\end{align*}
\]
We note that the master equation (11) includes only single-step processes where the transitions take place between the states \( n_1 \) and \( n_1 \pm 1 \), or \( n_2 \) and \( n_2 \pm 1 \). We can define new dimensionless quantities in terms of the reaction rates as follows:

\[
N_i = \frac{3\lambda_i}{2\tau_i}, \quad \delta_i^2 = 1 + \frac{8\tau_i(\mu_i - \gamma_i - d)}{3\lambda_i^2}, \quad R_{0}^{(i)} = \frac{\mu_i}{\gamma_i + d},
\]

(12)

Note that \( N_i \) defines the scale of the typical size of population \( i \) prior to extinction. The identities (12) establish a relation between the microscopic \((\lambda_i, \mu_i, \gamma_i, \tau_i, d)\) and macroscopic \((N_i, \delta_i, R_{0}^{(i)})\) parameters, which are obtainable through field observations. Note that the individual-based model displays the strong Allee effect, if \( \mu_i < \gamma_i + d \), or \( R_{0}^{(i)} < 1 \). Note that for \( R_{0}^{(i)} < 1 \) we have \( \delta_i < 1 \), where for the strong Allee effect, we must also demand that \( \delta_i > 0 \).

As mentioned in the Introduction, our starting point for the model used in this work is the deterministic skeleton for the non-dimensionalized model described by Johnson and Hastings, reproduced here for convenience.

\[
\dot{X}_1 = X_1(\beta_1 - (X_1 - 1)^2) + D(X_2 - X_1) \quad (13a)
\]
\[
\dot{X}_2 = X_2(\beta_2 - (X_2 - 1)^2) + D(X_1 - X_2) \quad (13b)
\]

In the model above, the parameter \( \beta_i \) represents a measure for the quality of the environment by the population denoted by \( X_i \). The parameter \( D \) denotes passive diffusion in the system and is an indicator of network connectivity. See the work by Johnson and Hastings for a detailed exposition of the model.

In order to make our subsequent analyses feasible, we aim to reduce the dimensionality of the parameter space for our stochastic model. So, for the sake of illustration, we treat the stochastic rate parameters in (9) as identical for both populations. This assumption appears to be reasonable in many cases, because of the presence of dispersal between populations in close proximity. Moreover, this assumption is consistent with the model parameterization in the work by Johnson and Hastings.

Thus, our matching scheme can be written as follows:

\[
\mu_1 := \beta_1, \quad (14a)
\]
\[
\mu_2 := \beta_2, \quad (14b)
\]
\[
\lambda := \lambda_1 = \lambda_2 \equiv 4, \quad (14c)
\]
\[
\gamma := \gamma_1 = \gamma_2 \equiv 1, \quad (14d)
\]
\[
\tau := \tau_1 = \tau_2 \equiv 6, \quad (14e)
\]
\[ d := D, \quad (14f) \]
\[ \tilde{N} := \tilde{N}_1 = \tilde{N}_2 \quad (14g) \]

So for \( i = 1 \) and 2,

\[ \tilde{N} = 1, \quad \delta_i^2 = \beta_i - D, \quad R_0^{(i)} = \frac{\beta_i}{D + 1}. \quad (15) \]

In what follows, we are interested in the case of bistability, which is manifest in the case of the strong Allee effect. Note that this necessitates the following condition:

\[ D < \beta_i < D + 1 \quad (16) \]

### 3 Methods

We note that our stochastic model operates over a two-dimensional state space. We argue that the dimension of the state space is as low as possible but nevertheless captures the desired phenomenology. Due to the presence of Allee effects, the deterministic skeleton of our model has cubic nonlinearities. In order to tackle this complexity, Johnson and Hastings conducted numerical simulations to understand how the transition rate \( D \) between the two patches determines the bifurcation structure for this system (Johnson and Hastings 2018). In this study, however, we restrict our attention to the case of bistability, which is properly addressed with a stochastic model.

As discussed in the previous section, we chose to adapt a master equation approach for our model (Méndez et al. 2019). Instead of specifying a model with carrying capacities, we used an individual-based modeling approach using a chemical reaction network. This allows for a fine-grained representation of the underlying discrete, stochastic process. Using this approach, we then wrote down the two-dimensional chemical master equation for the process. Given that this stochastic process is a continuous-time Markov chain (CTMC), it can be explicitly described by a generator \( Q \)-matrix with a countable state space. A nice feature of most ecological models is that they are built around processes that will approach a compact region exponentially quickly. Any reasonable ecological model should not have unbounded population growth. Density dependence in ecology models typically ensures this. Thus, our model effectively operates over a finite state space as the probability of arbitrarily large populations is negligibly small. Using this insight, we were able to obtain a finite-state CTMC in two dimensions.

Since the multi-dimensional master equation was relatively unwieldy to work with, we sought to reduce the two-dimensional state space to one dimension (Allen 2010; Allen and Allen 2003). Denoting \( N := N_1 = N_2 \) as the maximum number of individuals in either population, the specific mapping function used was \( f(x, y) = (N + 1)x + y + 1 \), for \( x, y \in [0, N] \). We could also exploit the sparsity of the banded \( Q \)-matrix to reduce computational cost (van Doorn and Pollett 2013). Specifically, the \( Q \)-matrix has size \((N + 1)^2 \times (N + 1)^2\) with \(7N^2 + 4N - 2\).
nonzero entries, yielding a matrix density of $O(N^{-2})$. Thus, the sparsity of the matrix increases quadratically in $N$.

For the ensuing analyses, we obtained population thresholds by treating them as identical across patches. Denoting $L := L_1 = L_2$ and $H := H_1 = H_2$ as the low and high thresholds, we specified $A := A_1 = A_2$ as the Allee threshold, noting that $0 < L < A < H \leq N$. Here, $N$ is defined as earlier, and can be understood as a system size parameter. The linear ordering of the thresholds guarantees that $L$ is always less than $H$. Note that the smallest $L$ threshold, $L = 1$, behaves as a quasi-extinct state. Moreover, with $2 \leq A \leq N - 1$, we have $A + 1 \leq H \leq N$ and $1 \leq L \leq A - 1$. It can be shown that this results in $(N^3)$ combinations of the thresholds.

In order to probe the system under consideration, we computed the mean first passage times (MFPTs) for model (9) with the state-space parameterized by $N$ for all combinations of the parameters $\beta_i, D, L,$ and $H$ (Chou and D’Orsogna 2014; Polizzi et al. 2016). First, we formed the $Q$-matrix for each point in parameter space. We also formed a vector of initial state probabilities $p_0$ governing the subsequent evolution of the CTMC. Then, the rows and columns of $Q$ corresponding to the trap states (e.g. extinction) were removed. Similarly, the corresponding entry in $p_0$ was removed. Next, using the truncated $Q$-matrix, $\tilde{Q}$, we computed the matrix of mean residence times, or $-\tilde{Q}^{-1}$. Finally, we computed the sum of the entries in $-\tilde{Q}^{-1} p_0$ to yield the MFPT from the initial state to the desired end state.

We could then construct a compartmental system with a reduced state space consisting of just four states: $HH$, $HL$, $LL$, and $LH$. Each of these states corresponds to a combination of population thresholds. For instance, $HL$ means that the first population is at a high abundance and the second population is at a low level. The MFPTs from the original model were used as input rates for the transition rate matrix of the reduced model. Thus, we used an emulator, or meta-model, as a proxy to analyze the original system.
Throughout this work, we analyze a simple system that exhibits noise-induced tipping point cascades in the vicinity of saddle-node bifurcations. We can use the emulator described previously to construct the schematic diagram in Fig. 1. In the diagram, each $r_i$ for $i = 1, \ldots, 4$ represents the rate of the transition between the relevant compartments. Each rate can be obtained as the inverse of the corresponding mean first passage time. Given that the process is represented as a CTMC, we can write down the transition rate matrix $S$ for the emulator. The transition probability matrix of the embedded discrete-time Markov chain gives the one-step transition probabilities of the system.

The $S$-matrix is given as:

$$S = \begin{pmatrix}
-r_1 - r_5 & r_1 & r_5 & 0 \\
r_8 & -r_2 - r_8 & 0 & r_2 \\
r_4 & 0 & -r_4 - r_6 & r_6 \\
0 & r_7 & r_3 & -r_3 - r_7
\end{pmatrix}$$

where the ordering of the states is $(HH, HL, LH, LL)$. Note that $S$ satisfies the properties of a generator matrix, namely:

- All off-diagonal elements are non-negative.
- All diagonal elements are negative.
- Each row sums to zero.

Due to the symmetry of the system, we have that $r_1 = r_5$, $r_2 = r_6$, $r_3 = r_7$, and $r_4 = r_8$. Thus

$$S = \begin{pmatrix}
-2r_1 & r_1 & r_1 & 0 \\
r_4 & -r_2 - r_4 & 0 & r_2 \\
r_4 & 0 & -r_2 - r_4 & r_2 \\
0 & r_3 & r_3 & -2r_3
\end{pmatrix}$$

Since the two populations $X_1$ and $X_2$ interact with each other via the symmetric dispersal parameter $D$ and the thresholds for both populations are equivalent, we can treat the states $LH$ and $HL$ as equivalent. Hence, we can simplify our formalism by omitting the $HL$ state (Fig. 2):

$$\tilde{S} = \begin{pmatrix}
-2r_1 & 2r_1 & 0 \\
2r_4 & -2r_2 - 2r_4 & 2r_2 \\
0 & 2r_3 & -2r_3
\end{pmatrix}$$
where the ordering of the states is now \((HH, LH, LL)\). Note that some of the rates are now doubled, since we have mapped two system states \((HL\) and \(LH)\) into one aggregate state \(LH\).

Now, we obtain the transition probability matrix \(\tilde{T}\) corresponding to the system:

\[
\tilde{T} = \begin{pmatrix}
0 & 1 & 0 \\
\frac{r_4}{r_2 + r_4} & 0 & \frac{r_2}{r_2 + r_4} \\
0 & 1 & 0
\end{pmatrix}
\]

Note that \(\tilde{T}\) is a right stochastic Markov matrix. Hence, the probability of a system collapse differs from the likelihood of system resiliency in a fundamental manner that is quantified by the parameters \(r_2 = \frac{1}{\text{MFPT}(LH \to LL)}\) and \(r_4 = \frac{1}{\text{MFPT}(LH \to HH)}\). In other words,

\[
P(LH \to HH) = \frac{r_4}{r_2 + r_4}
\]

denotes the likelihood of the system recovering to high population abundances, and

\[
P(LH \to LL) = \frac{r_2}{r_2 + r_4}
\]

represents the probability of the system collapsing. If \(r_2 = r_4\), then both probabilities are equal to \(1/2\), so the system is equally likely to collapse or recover.

Our aim henceforth is to explore the two-dimensional region of parameter space that corresponds to the desired probability of system resiliency. In symbols, for a fixed \(N\), we want to identify all combinations \((D, \beta_1, \beta_2, A, H, L)\) where \(D > 0\), \(\beta_1 \in (D, D + 1)\), \(\beta_2 \in (D, D + 1)\), \(A \in [2, N - 1]\), \(H \in [A + 1, N]\), and \(L \in [1, A - 1]\), such that the system recovers to a high-high state. In order to explore the parameter space numerically, we began by specifying an upper bound of 1 for the dispersal parameter \(D\). This value for \(D\) could be considered as large, since \(D\) should be commensurate with the per-capita rate of the system, which was chosen as \(r = 1\) in the non-dimensionalization of the model \((13a)-(13b)\) (Johnson and Hastings 2018). This implies that \(D \in (0, 1)\). Then, we chose \(N = 10\). This was the largest \(N\) such that the largest condition number among the space of all matrices \(\tilde{Q}\), taken over the set of simulated parameter values, was less than \(10^7\). For our chosen value of \(N\), we found this condition number to be approximately \(8.56 \times 10^6\). We note that numerical linear algebra was used here instead of Monte-Carlo methods, including the Doob-Gillespie algorithm (Gillespie 1976, 1977)). The primary justification for this choice of method was that some of the simulated mean first passage times were very large, causing prohibitively long runtimes with Monte-Carlo simulations. All computations were performed in MATLAB (MATLAB 2020).
Fig. 3 System resilience as a function of model parameters. A matrix of heatmaps summarizing the simulation study. Each heatmap indicates the value of $\nu$ corresponding to a parameter combination $(\beta_1, \beta_2, D, \eta)$. The range of $\beta_1$ values are on the x-axis and the range of values for $\beta_2$ are on the y-axis of each heatmap. In the $3 \times 3$ matrix comprising the entire figure, $D$ increases by row and $\eta$ increases by column. Since $0 \leq \nu \leq 1$, the colorbar is identical for each individual heatmap and ranges from 0 to 1 (Color figure online)

4 Results

By defining $r$ as the probability of transitioning from the $LH$ state to the $HH$ state, we can explore how the behavior of the system depends on the parameters. For purposes of display, we chose three values for the parameters governing patch quality, for each value of $D$ (i.e., $\beta_1 = D + 0.01$, $D + 0.5$, $D + 0.99$). By analyzing the simulation output, we found that the minimum value of $r$ over the parameter space was approximately 0.35, and the maximum value was 1. This indicates that it is possible to guarantee system recovery for some set of parameter values, but there is no set of parameter values that ensures system failure. In other words, the system is never more than 65% likely to fail. We also varied a threshold $\eta$, such that $r > \eta$, with sampling values of $\eta = 0.9, 0.95,$ and 0.99. We can expect that the percentage of parameter space $\nu$ for which the system recovers will decrease monotonically with the threshold $\eta$. If the goal is population persistence, we would like to find a region of parameter space that yields large values for both $\eta$ and $\nu$.

Referring to Fig. 3, several observations can be made. For instance, low dispersal ($D = 0.01$) guarantees that the desired probability of system resiliency is never 90% or higher (i.e., $\nu = 0$). This makes sense because rescue effects, by which the stronger patch saves the weaker patch, are reduced at low dispersal rates. Also, $\nu$ decreases as $\eta$ is increased. This is plausible since $\eta$ acts as a constraint on the parameter space and increasing $\eta$ implies greater specificity. In all cases, if both Allee effects are weakly strong (i.e., $(\beta_1, \beta_2) = (D + 0.99, D + 0.99)$), then $\nu$ is maximized for a given combination of $D$ and $\eta$. This is especially noticeable when $D = 0.99$, regardless of the value of the threshold parameter $\eta$ (i.e., see the last row of Fig. 3). Lastly, if each heatmap is treated as a matrix, we see that the column (row) of a given matrix
Fig. 4 Likelihood of positive tipping in the presence of low dispersal, with $L = 1$. A visualization of the case of low dispersal with a low threshold ($L$) corresponding to quasi-extinction. The combinations of $A$ and $H$ are shown explicitly. Each heatmap shows the value of $r$ corresponding to a combination $(\beta_1, \beta_2, H, A)$ for $D = 0.01$ and $L = 1$. In the matrix comprising the figure, $\beta_1$ increases by row and $\beta_2$ increases by column. Since $r$ is a probability, the colorbar is identical for each individual heatmap and ranges from 0 to 1 (Color figure online)

is increasing in $\nu$, as the corresponding column (row) indices are increased. This observation is also reflected within the coloring scheme of each matrix. As expected, $\nu$ is maximal with a value of 0.99, for $D = 0.99, \eta = 0.9, \beta_1 = D + 0.99, \text{and} \beta_2 = D + 0.99$.

We made explicit the correspondence between various combinations of the high and Allee thresholds, with $L$ fixed at a value of 1 (Fig. 5). We focused here on the case of low dispersal (i.e., $D = 0.01$). We see that the combination $(A, H) = (2, 3)$ is systematically undesirable for the system, as it always results in the lowest value of $r$ for a given pair $(\beta_1, \beta_2)$. However, as the high threshold is increased while the Allee threshold is held constant, we see that $r$ increases monotonically. This is true for any value of the Allee threshold (i.e., with only $H$ increased). Moreover, $r$ does not seem to depend much on either $A$ or $H$ for any fixed combination of the other parameters in the model. At least one of the Allee effects needs to be weakly strong (i.e., $D + 0.99$) to see a marked increase in $r$. We explored (Fig. 5) the case of moderate dispersal, with $D = 0.5$. Here, we notice that the lowest value for $r$ is 0.51, which is attained for $(A, H) = (2, 3)$, as in Fig. 4. The key difference is that $r$ now has a moderate dependence on both $A$ and $H$ for any fixed combination $(\beta_1, \beta_2, D, L)$. At least one of the Allee effects needs to be moderately strong (i.e., $D + 0.50$) to see a clear improvement in system resilience, or an increase in $r$. This means that all of the subcases except $(\beta_1, \beta_2) = (D + 0.01, D + 0.01)$ are sensitive to $r$. Finally, in Fig. 6, we address the scenario of high dispersal ($D = 0.99$. It is clear that $(A, H) = (2, 3)$
Fig. 5 Likelihood of positive tipping in the presence of moderate dispersal, with $L = 1$. A visualization of the case of medium dispersal with a low threshold ($L$) corresponding to quasi-extinction. The combinations of $A$ and $H$ are shown explicitly. Each heatmap shows the value of $r$ corresponding to a combination ($\beta_1, \beta_2, H, A$) for $D = 0.5$ and $L = 1$. In the matrix comprising the figure, $\beta_1$ increases by row and $\beta_2$ increases by column. Since $r$ is a probability, the colorbar is identical for each individual heatmap and ranges from 0 to 1 (Color figure online)

is a consistent minimizer across all levels of dispersal covered thus far, as the lowest value for $r$ is achieved there with a value of 0.82. Also, a vast majority of scenarios in the presence of high dispersal guarantee system recovery. Thus, $r$ is high throughout the subcases shown in Fig. 6.

For the sake of completeness, we also explored the analogues of Figs. 4, 5 and 6 by fixing $H = N$ and varying $L$. As shown in Fig. 7 corresponding to the case of low dispersal, $(A, L) = (9, 8)$ is an undesirable combination of thresholds for the system, as it guarantees the lowest value of $r$ for a given pair ($\beta_1, \beta_2$). Interestingly though, the highest value of $r$ occurs for $(\beta_1, \beta_2, A, L) = (0.99, 0.99, 9, 4)$, with the same maximal value of the Allee threshold as in the previous case. Furthermore, $r$ is maximized in the interior of the $A$–$L$ space, for any fixed combination of $\beta_1$ and $\beta_2$. Similarly, in Fig. 8 with moderate dispersal, the same combination of $(A, L) = (9, 8)$ yields the lowest value of $r$. On the other hand, $r$ is maximized at $(A, L) = (2, 1)$ instead. It happens that $r$ is maximized for very small and commensurate $A$ and $L$ in the case of moderate dispersal. Here, a vast majority of subcases favor system recovery. Finally, in Fig. 9, we analyze the case of high dispersal. Nearly all scenarios guarantee system recovery, with weaker Allee effects being more favorable for the system. Thus, $r$ has a weak dependence on parameters in the presence of high dispersal.

In general, a high degree of dispersal between both patches is needed to guarantee a desirable outcome for the system. In other words, in order to tip positive change with full confidence ($r \approx 1$), $D$ should be sufficiently large. A key issue which relates
The Role of Stochasticity in Noise-Induced Tipping Point… Page 15 of 20

Fig. 6 Likelihood of positive tipping in the presence of high dispersal, with $L = 1$. A visualization of the case of high dispersal with a low threshold ($L$) corresponding to quasi-extinction. The combinations of $A$ and $H$ are shown explicitly. Each heatmap shows the value of $r$ corresponding to a combination $(\beta_1, \beta_2, H, A)$ for $D = 0.99$ and $L = 1$. In the matrix comprising the figure, $\beta_1$ increases by row and $\beta_2$ increases by column. Since $r$ is a probability, the colorbar is identical for each individual heatmap and ranges from 0 to 1 (Color figure online)

to the idea of a rescue effect is the relative likelihood of the transition from $HH$ to $LH$, as compared with the transitions $LH$ to $LL$ and $LH$ to $HH$. The likelihood of system resiliency in the form of a rescue effect is given by $P(LH \rightarrow HH) = r$, whereas the probability of a total failure or catastrophic collapse is given by $P(LH \rightarrow LL) = 1 - r$. Note that this means a partial failure of the system is inevitable (i.e., with probability $1 = r + (1 - r)$). This result follows from the nature of the one-step transition probabilities in the emulator framework. Thus, we can reason that the odds of system recovery is equal to $\frac{r}{1-r}$.

Given this series of observations, there is an interesting and intuitive explanation that generalizes our setting. Noting that $0 \leq r \leq 1$ since $r$ is a probability, we have that

$$\frac{r}{1-r} = r + r^2 + r^3 + \ldots$$

is a convergent geometric series. Thus, we can write the left-hand side of the equation above as the odds of one network tipping favorably, where the network consists of a system with two patches. The right-hand side, of the equation, however, can be
Fig. 7 Likelihood of positive tipping in the presence of low dispersal, with $H = N$. A visualization of the case of low dispersal with a high threshold ($H$) corresponding to the carrying capacity. The combinations of $A$ and $L$ are shown explicitly. Each heatmap shows the value of $r$ corresponding to a combination ($\beta_1, \beta_2, A, L$) for $D = 0.01$ and $H = N$. In the matrix comprising the figure, $\beta_1$ increases by row and $\beta_2$ increases by column. Since $r$ is a probability, the colorbar is identical for each individual heatmap and ranges from 0 to 1 (Color figure online).

interpreted as the cumulative probability of all such networks tipping favorably, viz.

$$\frac{r}{1 - r} = \lim_{N \to \infty} \sum_{n=1}^{N} r^n,$$

where $n$ is the number of two-patch systems in a network. With this analysis, we thus see an instance of a tipping cascade of networks in the form of a domino effect. Note that the propagation of domino dynamics through a collection of networks requires strong connectivity, or high dispersal, in our model (Lenton 2020).

5 Discussion

Our analysis of a minimal, stochastic system that could have cascading tipping points yields some interesting insights. A novel result of this work is that noise-induced tipping can distinguish a catastrophic collapse and a successful recovery from the brink of extinction. This idea is expressed succinctly as the distinction between the tipping of positive change and the failure to adequately address an impending critical transition (Lenton 2020). In our two-population model, we showed that if the system is in the low/high state, a stochastic perturbation due to demographic noise results in
either a collapse of both populations (low/low state) or a full recovery to the high/high stable state. This may be a key feature of spatially connected populations with strong Allee effects, and analyzing these dynamics should inform management strategies for similar ecological systems. We found that a population with higher resilience, in the form of a larger population abundance, has the capacity to save its counterpart through a rescue effect. In particular, this occurs due to the presence of dispersal between the populations. In the presence of noise, patch homogeneity, and strong network connectivity, the system is more likely to exhibit a tipping point cascade (Lenton 2020).

Some discussion of systems with noise-induced tipping and their features is needed at this point. In situations where tipping one system increases the probability of another system tipping—for example, melting of the Greenland ice sheet increases the likelihood of failure of the Atlantic Meridional Overturning Circulation (AMOC) - the first system should act as a proxy for the entire network, in terms of a call to action. This is an important point, because by definition, our system can tip without warning, thus precluding detection through generic early-warning signals. However, we have been able to characterize the likelihood of its propensity to tip, either favorably or catastrophically (Lenton 2020). The results from our analysis support the hypothesis that both stochastic dynamics might play a key role in regime shifts and their associated tipping (Rocha et al. 2018). In particular, domino effects have relatively slow temporal dynamics and larger spatial scales; in our case, the spatial scale of the dispersal...
Fig. 9 Likelihood of positive tipping in the presence of high dispersal, with $H = N$. A visualization of the case of high dispersal with a high threshold ($H$) corresponding to the carrying capacity. The combinations of $A$ and $L$ are shown explicitly. Each heatmap shows the value of $r$ corresponding to a combination $(\beta_1, \beta_2, A, L)$ for $D = 0.99$ and $H = N$. In the matrix comprising the figure, $\beta_1$ increases by row and $\beta_2$ increases by column. Since $r$ is a probability, the colorbar is identical for each individual heatmap and ranges from 0 to 1 (Color figure online)

parameter dominates the time scale in magnitude. In addition, the one-step transition probabilities in our model support the notion that structural dependencies manifest as one-way interactions for the domino effect. This is a salient feature of regime shift couplings. In our system, we note that a partial collapse is inevitable. Most examples of regime shifts exhibiting domino effects concern the Earth system, including monsoon weakening and thermohaline circulation collapse, as well as nutrient transport mechanisms (Scheffer 2009; Rocha et al. 2018).

Our analysis is based on an individual-based model that describes the dynamics of the system under consideration and this approach should have general applicability. The mean-field description of this model necessarily includes a quadratic term. In other work (Abraham 1991; Klose et al. 2020), the proposed tipping element omits the quadratic term, in an attempt to describe a dangerous bifurcation in the form of a cusp catastrophe. This would correspond to zero mortality due to natural causes in our model (i.e., $\gamma = 0$). Similarly, the absence of the quadratic term would indicate that both the carrying capacity and Allee threshold for either population are identically zero (Johnson and Hastings 2018). Thus, we see that important details of the underlying mechanisms of a tipping element can be lost in a phenomenological description.

Ways to generalize our work include looking at more than two populations which would allow the study of how specific forms of connectivity could play a role in tipping cascades, or introducing the possibility of more complex dynamics (Strogatz 2001). The novelty of the emulator framework introduced here would be unaffected by the
inclusion of additional dimensions, or patches, since the required (but modified) mapping procedure to store matrix elements in memory would remain valid. Allowing for patch heterogeneity in the two populations by distinguishing their stochastic reaction rates could also produce interesting dynamics. We leave the exploration of these aspects to future work. Given the ubiquity of Allee effects (Courchamp et al. 2008) and the generality of our model, we believe that we have uncovered important ecological conclusions that are robust and should apply in a variety of settings. Within the field of landscape ecology, rescue effects are important in both metapopulations and species augmentation efforts. Another potential area of application is network theory.

Acknowledgements The authors would like to thank the anonymous referee for extensive and valuable feedback that improved the manuscript.

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

Abraham R (1991) Cuspoidal nets. In: Thygesen N et al (eds) Business cycles. Springer, Berlin, pp 56–63
Allen LJS (2010) An introduction to stochastic processes with applications to biology. CRC Press, Boca Raton
Allen LJS, Allen EJ (2003) A comparison of three different stochastic population models with regard to persistence time. Theor Popul Biol 64(4):439–449
Boukal DS, Berec L (2002) Single-species models of the allee effect: extinction boundaries, sex ratios and mate encounters. J Theor Biol 218(3):375–394
Chou T, D’Orsogna MR (2014) First passage problems in biology. In: Metzler R, Oshanin G, Redner S (eds) First-passage phenomena and their applications. World Scientific, Singapore, pp 306–345
Courchamp F, Clutton-Brock T, Grenfell B (1999) Inverse density dependence and the allee effect. Trends Ecol Evol 14(10):405–410
Courchamp F, Berec L, Gascoigne J (2008) Allee effects in ecology and conservation. Oxford University Press, Oxford
Dai L, Vorselen D, Korolev KS, Gore J (2012) Generic indicators for loss of resilience before a tipping point leading to population collapse. Science 336(6085):1175–1177
Drake JM, Kramer AM (2011) Allee effects. Nat Educ Knowl 3:2
Gardiner CW (2004) Handbook of stochastic methods: for physics, chemistry and the natural sciences. Springer, Berlin
Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22(4):403–434
Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81(25):2340–2361
Hastings A, Gross L (2012) Encyclopedia of theoretical ecology. University of California Press, Berkeley
Holling CS (1973) Resilience and stability of ecological systems. Annu Rev Ecol Syst 4(1):1–23
Johnson CL, Hastings A (2018) Resilience in a two-population system: interactions between allee effects and connectivity. Theor Ecol 11(3):281–289
Klose AK, Karle V, Winkelmann R, Donges JF (2020) Emergence of cascading dynamics in interacting tipping elements of ecology and climate. R Soc Open Sci 7(6):
Lenton TM (2020) Tipping positive change. Philos Trans R Soc B 375(1794):20190123
MATLAB (2020) R2020a. The MathWorks Inc., Natick
Méndez V, Assaf M, Masó-Puigdollerosa A, Campos D, Horstemkne W (2019) Demographic stochasticity and extinction in populations with allee effect. Phys Rev E 99(2):
Odum HT, Allee WC (1954) A note on the stable point of populations showing both intraspecific cooperation and disoperation. Ecology 35(1):95–97
O’Regan SM (2018) How noise and coupling influence leading indicators of population extinction in a spatially extended ecological system. J Biol Dyn 12(1):211–241
Polizzi NF, Therien MJ, Beratan DN (2016) Mean first-passage times in biology. Isr J Chem 56(9–10):816–824
Rocha JC, Peterson G, Bodin Ö, Levin S (2018) Cascading regime shifts within and across scales. Science 362(6421):1379–1383
Scheffer M (2009) Critical transitions in nature and society, vol 16. Princeton University Press, Princeton
Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock W, Dakos V, Van de Koppel J, Van de Leemput IA, Levin SA, Van Nes EH et al (2012) Anticipating critical transitions. Science 338(6105):344–348
Strogatz S (2001) Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering (studies in nonlinearity). Westview Press, Boulder
van Doorn EA, Pollett PK (2013) Quasi-stationary distributions for discrete-state models. Eur J Oper Res 230(1):1–14
Van Kampen NG (1992) Stochastic processes in physics and chemistry. Elsevier, Amsterdam
Volterra V (1938) Population growth, equilibria, and extinction under specified breeding conditions: a development and extension of the theory of the logistic curve. Hum Biol 10(1):1–11
Vortkamp I, Schreiber SJ, Hastings A, Hilker FM (2020) Multiple attractors and long transients in spatially structured populations with an allee effect. Bull Math Biol 82(6):1–21

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.