

Persistence in fluctuating environments for interacting structured populations

Gregory Roth · Sebastian J. Schreiber

Received: 23 May 2013 / Revised: 28 October 2013 / Published online: 6 December 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Individuals within any species exhibit differences in size, developmental state, or spatial location. These differences coupled with environmental fluctuations in demographic rates can have subtle effects on population persistence and species coexistence. To understand these effects, we provide a general theory for coexistence of structured, interacting species living in a stochastic environment. The theory is applicable to nonlinear, multi species matrix models with stochastically varying parameters. The theory relies on long-term growth rates of species corresponding to the dominant Lyapunov exponents of random matrix products. Our coexistence criterion requires that a convex combination of these long-term growth rates is positive with probability one whenever one or more species are at low density. When this condition holds, the community is stochastically persistent: the fraction of time that a species density goes below $\delta > 0$ approaches zero as δ approaches zero. Applications to predator-prey interactions in an autocorrelated environment, a stochastic LPA model, and spatial lottery models are provided. These applications demonstrate that positive autocorrelations in temporal fluctuations can disrupt predator-prey coexistence, fluctuations in log-fecundity can facilitate persistence in structured populations, and long-lived, relatively sedentary competing populations are likely to coexist in spatially and temporally heterogenous environments.

Mathematics Subject Classification 37H10 · 60G10 · 92D25 · 92D40

G. Roth (✉) · S. J. Schreiber
Department of Evolution and Ecology, One Shields Avenue, University of California,
Davis, CA 95616, USA
e-mail: greg.roth51283@gmail.com

S. J. Schreiber
e-mail: sschreiber@ucdavis.edu

1 Introduction

All populations are structured and experience environmental fluctuations. Population structure may arise to individual differences in age, size, and spatial location (Metz and Diekmann 1986; Caswell 2001; Holyoak et al. 2005). Temporal fluctuations in environmental factors such light, precipitation, and temperature occur in all natural marine, freshwater and terrestrial systems. Since these environmental factors can influence survival, growth, and reproduction, environmental fluctuations result in demographic fluctuations that may influence species persistence and the composition of ecological communities (Tuljapurkar 1990; Chesson 2000b; Kuang and Chesson 2009). Here we present, for the first time, a general approach to studying coexistence of structured populations in fluctuating environments.

For species interacting in an ecosystem, a fundamental question is what are the minimal conditions to ensure the long-term persistence of all species. Historically, theoretical ecologists characterize persistence by the existence of an asymptotic equilibrium in which the proportion of each population is strictly positive (May 1975; Roughgarden 1979). More recently, coexistence was equated with the existence of an attractor bounded away from extinction (Hastings 1988), a definition that ensures populations will persist despite small, random perturbations of the populations (Schreiber 2006, 2007). However, “environmental perturbations are often vigorous shake-ups, rather than gentle stirrings” (Jansen and Sigmund 1998). To account for large, but rare, perturbations, the concept of permanence, or uniform persistence, was introduced in late 1970s (Freedman and Waltman 1977; Schuster et al. 1979). Uniform persistence requires that asymptotically species densities remain uniformly bounded away from extinction. In addition, permanence requires that the system is dissipative i.e. asymptotically species densities remain uniformly bounded from above. Various mathematical approaches exist for verifying permanence (Hutson and Schmitt 1992; Smith and Thieme 2011) including topological characterizations with respect to chain recurrence (Butler and Waltman 1986; Hofbauer and So 1989), average Lyapunov functions (Hofbauer 1981; Hutson 1984; Garay and Hofbauer 2003), and measure theoretic approaches (Schreiber 2000; Hofbauer and Schreiber 2010). The latter two approaches involve the long-term, per-capita growth rates of species when rare. For discrete-time, unstructured models of the form $x_{t+1}^i = f_i(x_t)x_t^i$ where $x_t = (x_t^1, \dots, x_t^n)$ is the vector of population densities at time t , the long-term growth rate of species i with initial community state $x_0 = x$ equals

$$r_i(x) = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{s=0}^{t-1} \log f_i(x_s).$$

Garay and Hofbauer (2003) showed, under appropriate assumptions, that the system is permanent provided there exist positive weights p_1, \dots, p_n associated with each species such that $\sum_i p_i r_i(x) > 0$ for any initial condition x with one or more missing species (i.e. $\prod_i x^i = 0$). Intuitively, the community persists if on average the community increases when rare.

The permanence criterion for unstructured populations also extends to structured populations. However, in this case, the long-term growth rate is more complicated. Consider, for example, when both time and the structuring variables are discrete; the population dynamics are given by $x_{t+1}^i = A_i(x_t)x_t^i$ where x_t^i is a vector corresponding to the densities of the stages of species i , $x_t = (x_t^1, \dots, x_t^n)$, and $A_i(x)$ are non-negative matrices. Then the long term growth rate $r_i(x)$ of species i corresponds to the dominant Lyapunov exponent associated with the matrices $A_i(x)$ along the population trajectory:

$$r_i(x) = \limsup_{t \rightarrow \infty} \frac{1}{t} \log \|A_i(x_{t-1}) \dots A_i(x_0)\|.$$

At the extinction state $x = 0$, the long-term growth rate $r_i(0)$ simply corresponds to the log of the largest eigenvalue of $A_i(0)$. For structured single-species models, [Cushing \(1998\)](#) and [Kon et al. \(2004\)](#) proved that $r_1(0) > 0$ implies permanence. For structured, continuous-time, multiple species models, $r_i(x)$ can be defined in an analogous manner to the discrete-time case using the fundamental matrix of the variational equation. [Hofbauer and Schreiber \(2010\)](#) showed, under appropriate assumptions, that $\sum_i p_i r_i(x) > 0$ for all x in the extinction set is sufficient for permanence. For discrete-time structured models, however, there exists no general proof of this fact (see, however, [Salceanu and Smith 2009a,b, 2010](#)). When both time and the structuring variables are continuous, the models become infinite dimensional and may be formulated as partial differential equations or functional differential equations. Much work has been done in this direction ([Hutson and Moran 1987](#); [Zhao and Hutson 1994](#); [Thieme 2009, 2011](#); [Magal et al. 2010](#); [Xu and Zhao 2003](#); [Jin and Zhao 2009](#)). In particular, for reaction-diffusion equations, the long-term growth rates correspond to growth rates of semi-groups of linear operators and, $\sum_i p_i r_i(x) > 0$ for all x in the extinction set also ensures permanence for these models ([Hutson and Moran 1987](#); [Zhao and Hutson 1994](#); [Cantrell and Cosner 2003](#)).

Environmental stochasticity can be a potent force for disrupting population persistence yet maintaining biodiversity. Classical stochastic demography theory for stochastic matrix models $x_{t+1} = A(t)x_t$ shows that temporally uncorrelated fluctuations in the projection matrices $A(t)$ reduce the long-term growth rates of populations when rare ([Tuljapurkar 1990](#); [Boyce et al. 2006](#)). Hence, increases in the magnitude of these uncorrelated fluctuations can shift populations from persisting to asymptotic extinction. Under suitable conditions, the long-term growth rate for these models is given by the limit $r = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \|A(t) \dots A(1)\|$ with probability one. When $r > 0$, the population grows exponentially with probability one for these density-independent models. When $r < 0$, the population declines exponentially with probability one. [Hardin et al. \(1988\)](#) and [Benaim and Schreiber \(2009\)](#) proved that these conclusions extend to models with compensating density-dependence. However, instead of growing without bound when $r > 0$, the populations converge to a positive stationary distribution with probability one. These results, however, do not apply to models with over-compensating density-dependence or, more generally, non-monotonic responses of demography to density.

Environmental stochasticity can promote diversity through the storage effect (Chesson and Warner 1981; Chesson 1982) in which asynchronous fluctuations of favorable conditions can allow long-lived species competing for space to coexist. The theory for coexistence in stochastic environments has focused on stochastic difference equations of the form $x_{t+1}^i = x_t^i f_i(\xi_{t+1}, x_t)$ where ξ_1, ξ_2, \dots is a sequence of independent, identically distributed random variables (for a review see Schreiber 2012). Schreiber et al. (2011) prove that coexistence, in a suitable sense, occurs provided that $\sum_i p_i r_i(x) > 0$ with probability for all x in the extinction set. Similar to the deterministic case, the long-term growth rate of species i equals $r_i(x) = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{s=0}^{t-1} \log f_i(x_s)$. Here, stochastic coexistence implies that each species spends an arbitrarily small fraction of time near arbitrarily small densities.

Here, we develop persistence theory for models simultaneously accounting for species interactions, population structure, and environmental fluctuations. Our main result implies that the “community increases when rare” persistence criterion also applies to these models. Our model, assumptions, and a definition of stochastic persistence are presented in Sect. 2. Except for a compactness assumption, our assumptions are quite minimal allowing for overcompensating density dependence and correlated environmental fluctuations. Long-term growth rates for these models and our main theorem are stated in Sect. 3. We apply our results to stochastic models of predator-prey interactions, stage-structured beetle dynamics, and competition in spatial heterogeneous environments. The stochastic models for predator-prey interactions are presented in Sect. 4 and examine to what extent “colored” environmental fluctuations facilitate predator-prey coexistence. In Sect. 5, we develop precise criteria for persistence and exclusion for structured single species models and apply these results to the classic stochastic model of larvae-pupae-adult dynamics of flour beetles (Costantino et al. 1995; Dennis et al. 1995; Costantino et al. 1997; Henson and Cushing 1997) and metapopulation dynamics (Harrison and Quinn 1989; Gyllenberg et al. 1996; Metz and Gyllenberg 2001; Roy et al. 2005; Hastings and Botsford 2006; Schreiber 2010). We show, contrary to initial expectations, that multiplicative noise with logarithmic means of zero can facilitate persistence. In Sect. 6, we examine spatial-explicit lottery models (Chesson 1985, 2000a,b) to illustrate how spatial and temporal heterogeneity, collectively, mediate coexistence for transitive and intransitive competitive communities. Proofs of most results are presented in Sect. 8.

2 Model and assumptions

We study the dynamics of m interacting populations in a random environment. Each individual in population i can be in one of n_i individual states such as their age, size, or location. Let $X_t^i = (X_t^{i1}, \dots, X_t^{in_i})$ denote the row vector of populations abundances of individuals in different states for population i at time $t \in \mathbb{N}$. X_t^i lies in the non-negative cone $\mathbb{R}_+^{n_i}$. The *population state* is the row vector $X_t = (X_t^1, \dots, X_t^m)$ that lies in the non-negative cone \mathbb{R}_+^n where $n = \sum_{i=1}^m n_i$. To account for environment fluctuations, we consider a sequence of random variables, $\xi_1, \xi_2, \dots, \xi_t, \dots$ where ξ_t represents the state of the environment at time t .

To define the population dynamics, we consider projection matrices for each population that depend on the population state and the environmental state. More precisely, for each i , let $A_i(\xi, X)$ be a non-negative, $n_i \times n_i$ matrix whose j - k -th entry corresponds to the contribution of individuals in state j to individuals in state k e.g. individuals transitioning from state j to state k or the mean number of offspring in state k produced by individuals in state j . Using these projection matrices and the sequence of environmental states, the population dynamic of population i is given by

$$X_{t+1}^i = X_t^i A_i(\xi_{t+1}, X_t).$$

where X_t^i multiplies on the left hand side of $A_i(\xi_{t+1}, X_t)$ as it is a row vector. If we define $A(\xi, X)$ to be the $n \times n$ block diagonal matrix $\text{diag}(A_1(\xi, X), \dots, A_m(\xi, X))$, then the dynamics of the interacting populations are given by

$$X_{t+1} = X_t A(\xi_{t+1}, X_t). \tag{1}$$

For these dynamics, we make the following assumptions:

- H1: ξ_1, ξ_2, \dots is an ergodic stationary sequence in a compact Polish space E (i.e. compact, separable and completely metrizable).
- H2: For each i , $(\xi, X) \mapsto A_i(\xi, X)$ is a continuous map into the space of $n_i \times n_i$ non-negative matrices.
- H3: For each population i , the matrix A_i has fixed sign structure corresponding to a primitive matrix. More precisely, for each i , there is a $n_i \times n_i$, non-negative, primitive matrix P_i such that the j - k -th entry of $A_i(\xi, X)$ equals zero if and only if j - k -th entry P_i equals zero for all $1 \leq j, k \leq n_i$ and $(\xi, X) \in E \times \mathbb{R}_+^n$.
- H4: There exists a compact set $S \subset \mathbb{R}_+^n$ such that for all $X_0 \in \mathbb{R}_+^n$, $X_t \in S$ for all t sufficiently large.

Our analysis focuses on whether the interacting populations tend, in an appropriate stochastic sense, to be bounded away from extinction. Extinction of one or more population corresponds to the population state lying in the *extinction set*

$$S_0 = \left\{ x \in S : \prod_i \|x^i\| = 0 \right\}$$

where $\|x^i\| = \sum_{j=1}^{n_i} x^{ij}$ corresponds to the ℓ^1 -norm of x^i . Given $X_0 = x$, we define stochastic persistence in terms of the empirical measure

$$\Pi_t^x = \frac{1}{t} \sum_{s=1}^t \delta_{X_s} \tag{2}$$

where δ_y denotes a Dirac measure at y , i.e. $\delta_y(A) = 1$ if $y \in A$ and 0 otherwise for any Borel set $A \subset \mathbb{R}_+^n$. These empirical measures are random measures describing the distribution of the observed population dynamics up to time t . In particular, for any Borel set $B \subset S$,

$$\Pi_t^x(B) = \frac{\#\{1 \leq s \leq t | X_s \in B\}}{t}$$

is the fraction of time that the populations spent in the set B . For instance, if we define

$$S_\eta = \{x \in S : \|x^i\| \leq \eta \text{ for some } i\},$$

then $\Pi_t^x(S_\eta)$ is the fraction of time that the total abundance of some population is less than η given $X_0 = x$.

Definition 2.1 The model (1) is *stochastically persistent* if for all $\varepsilon > 0$, there exists $\eta > 0$ such that, with probability one,

$$\Pi_t^x(S_\eta) \leq \varepsilon$$

for t sufficiently large and $x \in S \setminus S_0$.

The set S_η corresponds to community states where one or more populations have a density less than η . Therefore, stochastic persistence corresponds to all populations spending an arbitrarily small fraction of time at arbitrarily low densities.

3 Results

3.1 Long-term growth rates and a persistence theorem

Understanding persistence often involves understanding what happens to each population when it is rare. To this end, we need to understand the propensity of the population to increase or decrease in the long term. Since

$$X_t^i = X_0^i A_i(\xi_1, X_0) A_i(\xi_2, X_1) \dots A_i(\xi_t, X_{t-1}),$$

one might be interested in the long-term “growth” of random product of matrices

$$A_i(\xi_1, X_0) A_i(\xi_2, X_1) \dots A_i(\xi_t, X_{t-1}) \tag{3}$$

as $t \rightarrow \infty$. One measurement of this long-term growth rate when $X_0 = x$ is the random variable

$$r_i(x) = \limsup_{t \rightarrow \infty} \frac{1}{t} \log \|A_i(\xi_1, X_0) A_i(\xi_2, X_1) \dots A_i(\xi_t, X_{t-1})\|. \tag{4}$$

Population i is tending to show periods of increase when $r_i(x) > 0$ and asymptotically decreasing when $r_i(x) < 0$. Since, in general, the sequence

$$\left\{ \frac{1}{t} \log \|A_i(\xi_1, X_0) A_i(\xi_2, X_1) \dots A_i(\xi_t, X_{t-1})\| \right\}_{t=1}^\infty$$

does not converge, the $\limsup_{t \rightarrow \infty}$ instead of $\lim_{t \rightarrow \infty}$ in the definition of $r_i(x)$ is necessary. However, as we discuss in Sect. 3.2, the $\limsup_{t \rightarrow \infty}$ can be replaced by $\lim_{t \rightarrow \infty}$ on sets of “full measure”.

An expected, yet useful property of $r_i(x)$ is that $r_i(x) \leq 0$ with probability one whenever $\|x^i\| > 0$. In words, whenever population i is present, its per-capita growth rate in the long-term is non-positive. This fact follows from X_t^i being bounded above for $t \geq 0$. Furthermore, on the event of $\{\limsup_{t \rightarrow \infty} \|X_t^i\| > 0\}$, we get that $r_i(x) = 0$ with probability one. In words, if population i 's density infinitely often is bounded below by some minimal density, then its long-term growth rate is zero as it is not tending to extinction and its densities are bounded from above. Both of these facts are consequences of results proved in the Sec. 8 (i.e. Proposition 8.10, Corollary 8.17 and Proposition 8.19).

Our main result extends the persistence conditions discussed in the introduction to stochastic models of interacting, structured populations. Namely, if the community increases on average when rare, then the community persists. More formally, we prove the following theorem in the Sec. 8.

Theorem 3.1 *If there exist positive constants p_1, \dots, p_m such that*

$$\sum_i p_i r_i(x) > 0 \quad \text{with probability one} \quad (5)$$

for all $x \in S_0$, then the model (1) is stochastically persistent.

For two competing species ($k = 2$) that persist in isolation (i.e. $r_1(0) > 0$ and $r_2(0) > 0$ with probability one), inequality (5) reduces to the classical mutual invasibility condition. To see why, consider a population state $x = (x_1, 0)$ supporting species 1. Since species 1 can persist in isolation, Proposition 8.19 implies that $r_1(x) = 0$ with probability one. Hence, inequality (5) for this initial condition becomes $p_1 r_1(x) + p_2 r_2(x) = p_2 r_2(x) > 0$ with probability one for all initial conditions $x = (x_1, 0)$ supporting species 1. Similarly, inequality (5) for an initial condition $x = (0, x_2)$ supporting species 2 becomes $r_1(x) > 0$ with probability one. In words, stochastic persistence occurs if both competitors have a positive per-capita growth rate when rare. A generalization of the mutual invasibility condition to higher dimensional communities is discussed at the end of the next subsection.

3.2 A refinement using invariant measures

The proof of Theorem 3.1 follows from a more general result that we now present. For this result, we show that one need not verify the persistence condition (5) for all x in the extinction set S_0 . It suffices to verify the persistence condition for invariant measures of the process supported by the extinction set.

Definition 3.2 A Borel probability measure μ on $E \times S$ is an invariant measure for the model (1) provided that

- (i) $\mathbb{P}[\xi_t \in B] = \mu(B \times S)$ for all Borel sets $B \subset E$, and
- (ii) if $\mathbb{P}[(\xi_0, X_0) \in C] = \mu(C)$ for all Borel sets $C \subset E \times S$, then $\mathbb{P}[(\xi_t, X_t) \in C] = \mu(C)$ for all Borel sets $C \subset E \times S$ and $t \geq 0$.

Condition (i) ensures that invariant measure is consistent with the environmental dynamics. Condition (ii) implies that if the system initially follows the distribution of μ , then it follows this distribution for all time. When this occurs, we say (ξ_t, X_t) is stationary with respect to μ . One can think of invariant measures as the stochastic analog of equilibria for deterministic dynamical systems; if the population statistics initially follow μ , then they follow μ for all time.

When an invariant measure μ is statistically indecomposable, it is ergodic. More precisely, μ is ergodic if it can not be written as a convex combination of two distinct invariant measures, i.e. if there exist $0 < \alpha < 1$ and two invariant measures μ_1, μ_2 such that $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$, then $\mu_1 = \mu_2 = \mu$.

Definition 3.3 If (ξ_t, X_t) is stationary with respect to μ , the subadditive ergodic theorem implies that $r_i(X_0)$ is well-defined with probability one. Moreover, we call the expected value

$$r_i(\mu) = \int \mathbb{E}[r_i(X_0)|X_0 = x, \xi_1 = \xi] \mu(d\xi, dx)$$

to be long-term growth rate of species i with respect to μ . When μ is ergodic, the subadditive ergodic theorem implies that $r_i(X_0)$ equals $r_i(\mu)$ for μ -almost every (X_0, ξ_1) .

With these definitions, we can rephrase Theorem 3.1 in terms of the long-term growth rates $r_i(\mu)$ as well as provide an alternative characterization of the persistence condition.

Theorem 3.4 *If one of the following equivalent conditions hold*

- (i) $r_*(\mu) := \max_{1 \leq i \leq m} r_i(\mu) > 0$ for every invariant probability measure with $\mu(S_0) = 1$, or
- (ii) there exist positive constants p_1, \dots, p_m such that

$$\sum_i p_i r_i(\mu) > 0$$

for every ergodic probability measure with $\mu(S_0) = 1$, or

- (iii) there exist positive constants p_1, \dots, p_m such that

$$\sum_i p_i r_i(x) > 0 \quad \text{with probability one}$$

for all $x \in S_0$

then the model (1) is stochastically persistent.

With Theorem 3.4's formulation of the stochastic persistence criterion, we can introduce a generalization of the mutual invasibility condition to higher-dimensional communities. To state this condition, observe that for any ergodic, invariant measure μ , there is a unique set of species $I \subset \{1, \dots, k\}$ such that $\mu(\{x : x^{ij} > 0 \text{ for all } i \in I, 1 \leq j \leq n_i\}) = 1$. In other words, μ supports the community I . Proposition 8.19 implies that $r_i(\mu) = 0$ for all $i \in I$. Therefore, if I is a strict subset of $\{1, \dots, k\}$ i.e. not all species are in the community I , then coexistence condition (ii) of Theorem 3.4 requires that there exists a species $i \notin I$ such that $r_i(\mu) > 0$. In other words, the coexistence condition requires that at least one missing species has a positive per-capita growth rate for any subcommunity represented by an ergodic invariant measure. While this weaker condition is sometimes sufficient to ensure coexistence (e.g. in the two species models that we examine), in general it is not as illustrated in Sect. 6.1. Determining, in general, when this "at least one missing species can invade" criterion is sufficient for stochastic persistence is an open problem.

4 Predator-prey dynamics in auto-correlated environments

To illustrate the applicability of Theorems 3.1 and 3.4, we apply the persistence criteria to stochastic models of predator-prey interactions, stage-structured populations with over-compensating density-dependence, and transitive and intransitive competition in spatially heterogeneous environments.

For unstructured populations, Theorem 3.4 extends Schreiber et al. (2011)'s criteria for persistence to temporally correlated environments. These temporal correlations can have substantial consequences for coexistence as we illustrate now for a stochastic model of predator-prey interactions. In the absence of the predator, assume the prey, with density N_t at time t , exhibits a noisy Beverton-Holt dynamic

$$N_{t+1} = \frac{R_{t+1}N_t}{1 + aN_t} \quad (6)$$

where R_t is a stationary, ergodic sequence of random variables corresponding to the intrinsic fitness of the prey at time t , and $a > 0$ corresponds to the strength of intraspecific competition. To ensure the persistence of the prey in the absence of the predator, assume $\mathbb{E}[\ln R_1] > 0$ and $\mathbb{E}[\ln R_1] < \infty$. Under these assumptions, Theorem 1 of Benaïm and Schreiber (2009) implies that N_t converges in distribution to a positive random variable \widehat{N} whenever $N_0 > 0$. Moreover, the empirical measures $\Pi_t^{(N,P)}$ with $N > 0$, $P = 0$ converge almost surely to the law ν of the random vector $(\widehat{N}, 0)$ i.e. the probability measure satisfying $\nu(A) = \mathbb{P}[(\widehat{N}, 0) \in A]$ for any Borel set $A \subset \mathbb{R}_+^2$.

Let P_t be the density of predators at time t and $\exp(-bP_t)$ be the fraction of prey that "escape" predation during generation t where b is the predator attack rate. The mean number of predators offspring produced per consumed prey is c , while s corresponds to the fraction of predators that survive to the next time step. The predator-prey dynamics are

$$N_{t+1} = \frac{R_{t+1}N_t}{1 + aN_t} \exp(-bP_t)P_{t+1} = cN_t(1 - \exp(-bP_t)) + sP_t. \quad (7)$$

To see that (7) is of the form of our models (1), we can expand the exponential term in the second equation. To ensure that (7) satisfies the assumptions of Theorem 3.1, we assume R_t takes values in the half open interval $(0, R^*]$. Since $N_{t+1} \leq R_{t+1}/a \leq R^*/a$ and $P_{t+1} \leq cN_t + sP_t \leq cR^*/a + sP_t$, $X_t = (N_t, P_t)$ eventually enters and remains in the compact set

$$S = [0, R^*/a] \times [0, cR^*/(a(1 - s))].$$

To apply Theorem 3.1, we need to evaluate $r_i((N, P))$ for all $N \geq 0, P \geq 0$ with either $N = 0$ or $P = 0$. Since $(0, P_t)$ converges to $(0, 0)$ with probability one whenever $P_0 \geq 0$, we have $r_1((0, P)) = \mathbb{E}[\ln R_t] > 0$ and $r_2((0, P)) = \ln s < 0$ whenever $P \geq 0$. Since $\Pi_t^{(N,0)}$ with $N > 0$ converges almost surely to ν , Proposition 8.19 implies $r_1((N, 0)) = 0$. Moreover,

$$\begin{aligned} r_2((N, 0)) &= \mathbb{E} [\ln (cb\widehat{N} + s)] \\ &= \int \ln (cbx + s)\nu(dx). \end{aligned} \tag{8}$$

By choosing $p_1 = 1 - \varepsilon$ and $p_2 = \varepsilon > 0$ for ε sufficiently small (e.g. $0.5\mathbb{E}[\ln R_t]/(\mathbb{E}[\ln R_t] - \ln s)$), we have $\sum_i p_i r_i((N, P)) > 0$ whenever $NP = 0$ if and only if

$$\mathbb{E} [\ln (cb\widehat{N} + s)] > 0. \tag{9}$$

Namely, the predator and prey coexist whenever the predator can invade the prey-only system. Since $\ln(cbN + s)$ is a concave function of the prey density and the predator life history parameters c, b, s , Jensen’s inequality implies that fluctuations in any one of these quantities decreases the predator’s growth rate.

To see how temporal correlations influence whether the persistence criterion (9) holds or not, consider an environment that fluctuates randomly between good and bad years for the prey. On good years, R_t takes on the value R_{good} , while in bad years it takes on the value R_{bad} . Let the transitions between good and bad years be determined by a Markov chain where the probability of going from a bad year to a good year is p and the probability of going from a good year to a bad year is q . For simplicity, we assume that $p = q$ in which case half of the years are good and half of the years are bad in the long run. Under these assumptions, the persistence assumption $\mathbb{E}[\ln R_1] > 0$ for the prey is $\ln (R_{good}R_{bad}) > 0$.

To estimate the left-hand side of (9), we consider the limiting cases of strongly negatively correlated environments ($p \approx 1$) and strongly positively correlated environments ($p \approx 0$). When $p \approx 1$, the environmental dynamics are nearly periodic switching nearly every other time step between good and bad years. Hence, one can approximate the stationary distribution \widehat{N} by the positive, globally stable fixed point of

$$\begin{aligned}
 x_{t+2} &= \frac{R_{good}x_{t+1}}{1 + ax_{t+1}} \\
 &= \frac{R_{good}R_{bad}x_t/(1 + ax_t)}{1 + a(R_{bad}x_t/(1 + ax_t))} \\
 &= \frac{R_{good}R_{bad}x_t}{1 + a(1 + R_{bad})x_t}
 \end{aligned}$$

which is given by $\frac{R_{good}R_{bad}-1}{a(1+R_{bad})}$. Hence, if $p \approx 1$, then the distribution ν of \widehat{N} approximately puts half of its weight on $\frac{R_{good}R_{bad}-1}{a(1+R_{bad})}$ and half of its weight on $\frac{R_{good}R_{bad}-1}{a(1+R_{good})}$ and the persistence criterion (9) is approximately

$$\frac{1}{2} \ln \left(bc \frac{R_{good}R_{bad} - 1}{a(1 + R_{bad})} + s \right) + \frac{1}{2} \ln \left(bc \frac{R_{good}R_{bad} - 1}{a(1 + R_{good})} + s \right) > 0. \tag{10}$$

Next, consider the case that $p \approx 0$ in which there are long runs of good years and long runs of bad years. Due to these long runs, one expects that half time \widehat{N} is near the value $(R_{good} - 1)/a$ and half the time it is near the value $\max\{(R_{bad} - 1)/a, 0\}$. If $R_{bad} > 1$, then the persistence criterion is approximately

$$\frac{1}{2} \ln \left(bc \frac{R_{good} - 1}{a} + s \right) + \frac{1}{2} \ln \left(bc \frac{R_{bad} - 1}{a} + s \right) > 0. \tag{11}$$

Relatively straightforward algebraic manipulations (e.g. exponentiating the left hand sides of (10) and (11) and multiplying by $(1 + R_{bad})(1 + R_{good})$) show that the left hand side of (10) is always greater than the left hand side of (11).

Biological Interpretation 4.1 *Positive autocorrelations, by increasing variability in prey density, hinders predator establishment and, thereby, coexistence of the predator and prey. In contrast, negative auto-correlations by reducing variability in prey density can facilitate predator-prey coexistence (Fig. 1).*

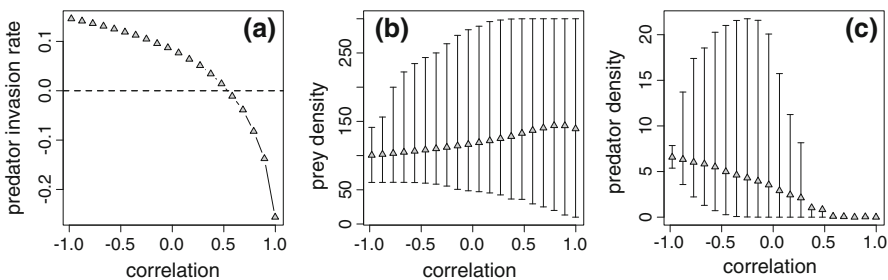


Fig. 1 Effect of temporal autocorrelations on predator-prey coexistence in a Markovian environment. **a** The long-term growth rate $r_2(N, 0)$ with $N > 0$ of the predator when rare is plotted as a function of the temporal autocorrelation between good and bad reproductive years for the prey. **b, c** The mean and interquartile ranges of long-term distribution of prey and predator densities are plotted as function of the temporal autocorrelation. Parameters: $R_{good} = 4, R_{bad} = 1.1, a = 0.01, c = 1, s = 0.1, b = 0.01$

5 Application to structured single species models

For single species models with negative-density dependence, we can prove sufficient and necessary conditions for stochastic persistence. The following theorem implies that stochastic persistence occurs if the long-term growth rate $r_1(0)$ when rare is positive and asymptotic extinction occurs with probability one if this long-term growth rate is negative.

Theorem 5.1 *Assume that $m = 1$ (i.e. there is one species), **H1-H4** hold and the entries of $A(\xi, x) = A_1(\xi, x)$ are non-increasing functions of x . If $r_1(0) > 0$, then*

$$X_{t+1} = X_t A(\xi_{t+1}, X_t) \quad (12)$$

is stochastically persistent. If $r_1(0) < 0$, then $\lim_{t \rightarrow \infty} X_t = (0, 0, \dots, 0)$ with probability one.

Our assumption that the entries $A(\xi, x)$ are non-increasing functions of x ensures that $r_1(0) \geq r_1(x)$ for all x which is the key fact used in the proof of Theorem 5.1. It remains an open problem to identify other conditions on $A(\xi, x)$ that ensure $r_1(0) \geq r_1(x)$ for all x .

Proof The first statement of this theorem follows from Theorem 3.1.

Assume that $r_1(0) < 0$. Provided that X_0 is nonnegative with at least one strictly positive entry, Ruelle's stochastic version of the Perron Frobenius Theorem (Ruelle 1979b, Proposition 3.2) and the entries of $A(\xi, x)$ being non-increasing in x imply

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|X_t\| \leq \lim_{t \rightarrow \infty} \frac{1}{t} \log \|X_0 A(\xi_t, 0) \dots A(\xi_1, 0)\| = r(0) < 0$$

with probability one. Hence, $\lim_{t \rightarrow \infty} X_t = (0, \dots, 0)$ with probability one. \square

Theorem 5.1 extends Theorem 1 of Benaïm and Schreiber (2009) as it allows for over-compensating density dependence and makes no assumptions about differentiability of $x \mapsto A(\xi, x)$. To illustrate its utility, we apply this result to the larvae-pupae-adult model of flour beetles and a metapopulation model.

5.1 A stochastic Larvae-Pupae-Adult model for flour beetles

An important, empirically validated model in ecology is the ‘‘Larvae-Pupae-Adult’’ (LPA) model which describes flour beetle population dynamics (Costantino et al. 1995; Dennis et al. 1995; Costantino et al. 1997). The model keeps track of the densities ℓ_t, p_t, a_t of larvae, pupae, and adults at time t . Adults produce b eggs each time step. These eggs are cannibalized by adults and larvae at rates c_{ea} and c_{el} , respectively. The eggs escaping cannibalism become larvae. A fraction μ_l of larvae die at each time step. Larvae escaping mortality become pupae. Pupae are cannibalized by adults at a rate c_{pa} . Those individuals escaping cannibalism become adults. A fraction μ_a of adults

survive through a time step. These assumptions result in a system of three difference equations

$$\begin{aligned}\ell_{t+1} &= b a_t \exp(-c_{el} \ell_t - c_{ea} a_t) \\ p_{t+1} &= (1 - \mu_l) \ell_t \\ a_{t+1} &= (p_t \exp(-c_{pa} a_t) + (1 - \mu_a) a_t)\end{aligned}\tag{13}$$

Environmental fluctuations have been included in these models in at least two ways. [Dennis et al. \(1995\)](#) assumed that each stage experienced random fluctuations due to multiplicative factors $\exp(\xi_t^l)$, $\exp(\xi_t^p)$, $\exp(\xi_t^a)$ such that ξ_t^i for $i = l, p, a$ are independent and normally distributed with mean zero i.e. on the log-scale the average effect of environmental fluctuations are accounted for by the deterministic model. Alternatively, [Henson and Cushing \(1997\)](#) considered periodic fluctuations in cannibalism rates due to fluctuations in the size V_t of the habitat i.e. the volume of the flour. In particular, they assumed that $c_i = \kappa_i / V_t$ for $i = ea, el, pa$, for positive constants κ_i . If we include both of these stochastic effects into the deterministic model, we arrive at the following system of random difference equations:

$$\begin{aligned}\ell_{t+1} &= b a_t \exp(-\kappa_{el} \ell_t / V_{t+1} - \kappa_{ea} a_t / V_{t+1} + \xi_t^l) \\ p_{t+1} &= (1 - \mu_l) \ell_t \exp(\xi_t^p) \\ a_{t+1} &= (p_t \exp(-\kappa_{pa} a_t / V_{t+1}) + (1 - \mu_a) a_t) \exp(\xi_t^a).\end{aligned}\tag{14}$$

We can use [Theorem 3.4](#) to prove the following persistence result. In the case of $\xi_t^i = 0$ with probability one for $i = l, p, a$, this theorem can be viewed as a stochastic extension of [Theorem 4](#) of [Henson and Cushing \(1997\)](#) for periodic environments.

Theorem 5.2 *Assume $c_i > 0$ for $i = ea, el, pa$, $\mu_i \in (0, 1)$ for $i = l, a$, $\xi_t^l, \xi_t^p, \xi_t^a$, and V_t are ergodic and stationary sequences such that $\xi_t^i, \log V_t \in (-M, M)$ for $i = l, p, a, t \geq 0$ and some $M > 0$, and $(1 - \mu_a) \exp(\xi_t^a) \in [0, 1 - \delta]$ for some $\delta > 0$ with probability one. Then there exists a critical birth rate $b_{crit} > 0$ such that*

Extinction: If $b < b_{crit}$, then $X_t = (\ell_t, p_t, a_t)$ converges almost surely to $(0, 0, 0)$ as $t \rightarrow \infty$.

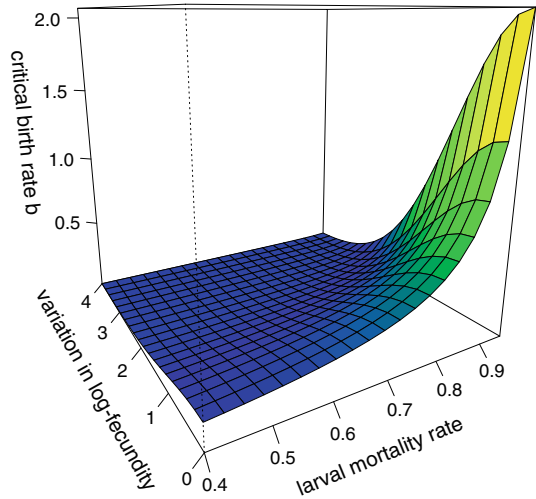
Stochastic persistence: If $b > b_{crit}$, then the LPA model is stochastically persistent.

Moreover, if $\xi_t^l = \xi_t^a = \xi_t^p$ with probability one and $\mathbb{E}[\xi_t^l] = 0$, then $b_{crit} = \mu_a / (1 - \mu_l)$.

Remark The assumption that ξ_t^i are compactly supported formally excludes the normal distributions used by [Dennis et al. \(1995\)](#). However, truncated normals with a very large M can approximate the normal distribution arbitrarily well. The assumption $(1 - \mu_a) \exp(\xi_t^a) \in [0, 1 - \delta]$ for some $\delta > 0$ is more restrictive. However, from a biological standpoint, it is necessary as this term corresponds to the fraction of adults surviving to the next time step. None the less, we conjecture that the conclusions of [Theorem 5.2](#) hold when ξ_t^i are normally distributed with mean 0.

[Theorem 5.2](#) implies that including multiplicative noise with log-mean zero has no effect on the deterministic persistence criterion when $\xi_t^l = \xi_t^p = \xi_t^a$ with probability

Fig. 2 Effects fluctuations in fecundity and larval survival on the critical birth rate b required for persistence. (ξ_t^l) are normally distributed with mean 0 and variance one, $\xi_t^a = \xi_t^p = 0$ for all t and $\mu_a = 0.1034$ (the value found in Table 1D in Costantino et al. (1995))



one. However, when these random variables are not perfectly correlated, we conjecture that this form of multiplicative noise always decreases the critical birth rate (Fig. 2). To provide some mathematical evidence for this conjecture, we compute a small noise approximation for the per-capita growth rate $r_1(\delta_0)$ when the population is rare (Ruelle 1979b; Tuljapurkar 1990). Let

$$B_t = \begin{pmatrix} 0 & (1 - \mu_l) \exp(\xi_t^p) & 0 \\ 0 & 0 & \exp(\xi_t^a) \\ b \exp(\xi_t^l) & 0 & (1 - \mu_a) \exp(\xi_t^a) \end{pmatrix}$$

be the linearization of the stochastic LPA model (14) at $(L, P, A) = (0, 0, 0)$. Assume that $\xi_t^i = \varepsilon Z_t^i$ where $\mathbb{E}[Z_t^i] = 0$ and $\mathbb{E}[(Z_t^i)^2] = 1$. Ruelle (1979b, Theorem 3.1) implies that $r_1(0)$ is an analytic function of ε . Therefore, one can perform a Taylor’s series expansion of $r_1(0)$ as function of ε about the point $\varepsilon = 0$. As we shall shortly show, the first non-zero term of this expansion is of second order. Expanding B_t to second order in ε yields

$$B_t \approx \underbrace{\begin{pmatrix} 0 & (1 - \mu_l) & 0 \\ 0 & 0 & 1 \\ b & 0 & (1 - \mu_a) \end{pmatrix}}_{=B} \left(I + \varepsilon \text{diag}\{Z_t^l, Z_t^p, Z_t^a\} + \varepsilon^2 \text{diag}\{Z_t^l, Z_t^p, Z_t^a\}^2 / 2 \right).$$

The entries of the second order term are positive due to the convexity of the exponential function. Hence, Jensen’s inequality implies that fluctuations in Z_t^i increase the mean matrix $\mathbb{E}[B_t]$. This observation, in and of itself, suggests that fluctuations in Z_t^i increase $r_1(0)$. However, to rigorously verify this assertion, let v and w be the left and right Perron-eigenvectors of B such that $\sum_i v_i = 1$ and $\sum_i v_i w_i = 1$. Let r_0 be the associated Perron eigenvalue of B . Provided the Z_t^i are independent in time, a small noise approximation for the stochastic growth rate of the random products of B_t is

$$r_*(\delta_0) \approx \log r_0 + \frac{\varepsilon^2}{2} \left(\mathbb{E} \left[\sum_i v_i w_i (Z_t^i)^2 \right] - \mathbb{E} \left[\left(\sum_i v_i w_i Z_t^i \right)^2 \right] \right). \tag{15}$$

Since the function $x \mapsto x^2$ is strictly convex and $\sum_i v_i w_i (Z_t^i)^2$ is a convex combination of $(Z_t^l)^2$, $(Z_t^p)^2$, and $(Z_t^a)^2$, Jensen’s inequality implies

$$\left(\sum_i v_i w_i Z_t^i \right)^2 \leq \sum_i v_i w_i (Z_t^i)^2.$$

Therefore

$$\mathbb{E} \left[\left(\sum_i v_i w_i Z_t^i \right)^2 \right] \leq \mathbb{E} \left[\sum_i v_i w_i (Z_t^i)^2 \right].$$

It follows that the order ε^2 correction term in (15) is non-negative and equals zero if and only if $Z_t^l = Z_t^p = Z_t^a$ with probability one. Therefore, “small” multiplicative noise (with log-mean zero) which isn’t perfectly correlated across the stages increases the stochastic growth rate and, therefore, decreases the critical birth rate b_{crit} required for stochastic persistence.

Biological Interpretation 5.3 *For the LPA model, there is a critical mean fecundity, above which the population persists and below which the population goes asymptotically to extinction. Fluctuations in the log survival rates decrease the critical mean fecundity unless the log survival rates are perfectly correlated.*

Proof of Theorem 5.2 We begin by verifying **H1–H4**. **H1** and **H2** follow from our assumptions. To verify **H3**, notice that the sign structure of the nonlinear projection matrix $A_t(\xi, X)$ for (14) is given by

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Since

$$C^4 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{pmatrix}$$

$A_t(\xi, X)$ has the sign structure of the primitive matrix C for all ξ, X and t . Finally, to verify **H4**, define

$$K = be^{2M-1}/\kappa_{ea}$$

Then

$$\ell_{t+1} \leq ba_t \exp(-\kappa_{ea}a_t/V_{t+1} + \xi_t^l) \leq ba_t \exp(-\kappa_{ea}a_t \exp(-M) + M) \leq K$$

for all $t \geq 0$. Therefore, $\ell_t \leq K$ for $t \geq 1$ and

$$p_t \leq \ell_{t-1}e^M \leq Ke^M$$

for all $t \geq 2$. Hence,

$$a_{t+1} \leq p_t e^M + (1 - \delta)a_t$$

for all $t \geq 2$ which implies $a_t \leq Ke^{3M}/\delta$ for t sufficiently large. The compact forward invariant set $S = [0, K] \times [0, Ke^M] \times [0, Ke^{3M}]/\delta$ satisfies **H4**.

At low density we get

$$B_t = A(\xi_t, 0) = \begin{pmatrix} 0 & (1 - \mu_l) \exp(\xi_t^p) & 0 \\ 0 & 0 & \exp(\xi_t^a) \\ b \exp(\xi_t^l) & 0 & (1 - \mu_a) \exp(\xi_t^a) \end{pmatrix}.$$

Define $r(b)$ to be the dominant Lyapunov exponent of the random products of B_1, B_2, \dots . Note that with the notation of Theorem 3.4, $r(b) = r_1(0)$. Theorem 3.1 of Ruelle (1979b) implies that $r(b)$ is differentiable for $b > 0$ and the derivative is given by (see, e.g., section 4.1 of Ruelle 1979b)

$$r'(b) = \mathbb{E} \left[\frac{v_t(b)E_{31}w_{t+1}(b)}{v_t(b)B_t(b)w_{t+1}(b)} \right] > 0$$

where $v_t(b), w_t(b)$ are the normalized left and right invariant sub-bundles associated with $B_t(b)$ and E_{31} is the matrix with $\exp(\xi_t^l)$ in the 3 – 1 entry and 0 entries otherwise. Since the numerator and denominators in the expectation are always positive, $r(b)$ is a strictly increasing function of b . Since $\lim_{b \rightarrow 0} r(b) = -\infty$ and $\lim_{b \rightarrow \infty} r(b) = \infty$, there exists $b_{crit} > 0$ such that $r(b) < 0$ for $b < b_{crit}$ and $r(b) > 0$ for $b > b_{crit}$.

If $b > b_{crit}$, then $r(b) > 0$ and Theorem 5.1 implies that (14) is stochastically persistent. On the other hand, if $b < b_{crit}$, then $r(b) < 0$ and Theorem 5.1 implies that (ℓ_t, p_t, a_t) converges to $(0, 0, 0)$ with probability one as $t \rightarrow \infty$.

The final assertion about the stochastic LPA model follows from observing that if $\xi_t^a = \xi_t^l = \xi_t^a$ with probability one for all t , then

$$B_t = \begin{pmatrix} 0 & (1 - \mu_l) & 0 \\ 0 & 0 & 1 \\ b & 0 & (1 - \mu_a) \end{pmatrix} \exp(\xi_t^l)$$

with probability one. Hence, $r(b) = \log r_0(b) + \mathbb{E}[\xi_t^l]$ where $r_0(b)$ is the dominant eigenvalue of the deterministic matrix

$$\begin{pmatrix} 0 & (1 - \mu_l) & 0 \\ 0 & 0 & 1 \\ b & 0 & (1 - \mu_a) \end{pmatrix}.$$

Therefore, if $\mathbb{E}[\xi_t^l] = 0$, then $r(b) = \log r_0(b)$. Using the Jury conditions, [Henson and Cushing \(1997\)](#) showed that $r_0(b) > 1$ if $b > \mu_a/(1 - \mu_l)$ and $r_0(b) < 1$ if $b < \mu_a/(1 - \mu_l)$. Hence, when $\xi_t^l = \xi_t^p = \xi_t^a$ with probability one and $\mathbb{E}[\xi_t^l] = 0$, b_{crit} equals $\mu_a/(1 - \mu_l)$ as claimed. \square

5.2 Metapopulation dynamics

Interactions between movement and spatio-temporal heterogeneities determine how quickly a population grows or declines. Understanding the precise nature of these interactive effects is a central issue in population biology receiving increasing attention from theoretical, empirical, and applied perspectives ([Petchey et al. 1997](#); [Lundberg et al. 2000](#); [Gonzalez and Holt 2002](#); [Schmidt 2004](#); [Roy et al. 2005](#); [Boyce et al. 2006](#); [Hastings and Botsford 2006](#); [Mathews and Gonzalez 2007](#); [Schreiber 2010](#)).

A basic model accounting for these interactions considers a population living in an environment with n patches. Let X_t^r be the number of individuals in patch r at time t . Assuming Ricker density-dependent feedbacks at the patch scale, the fitness of an individual in patch r is $\lambda_t^r \exp(-\alpha_r X_t^r)$ at time t , where λ_t^r is the maximal fitness and $\alpha_r > 0$ measures the strength of intraspecific competition. Let d_{rs} be the fraction of the population from patch r that disperse to patch s . Under these assumptions, the population dynamics are given by

$$X_{t+1}^r = \sum_{s=1}^n d_{sr} \lambda_t^s X_t^s \exp(-\alpha_s X_t^s) \quad r = 1, \dots, n. \tag{16}$$

To write this model more compactly, let $F(X_t, \lambda_t)$ be the diagonal matrix with diagonal entries $\lambda_1 \exp(-\alpha_1 X_t^1), \dots, \lambda_n \exp(-\alpha_n X_t^n)$, and D be the matrix whose i - j th entry is given by d_{ij} . With this notation, (16) simplifies to

$$X_{t+1} = X_t F(X_t, \lambda_t) D.$$

If λ_t^r are ergodic and stationary, λ_t^r take values in a positive compact interval $[\lambda_*, \lambda^*]$ and D is a primitive matrix, then the hypotheses of [Theorem 5.1](#) hold. In particular, stochastic persistence occurs only if $r_1(0)$, corresponding to the dominant Lyapunov exponent of the random matrix product $F(0, \lambda_t) D \dots F(0, \lambda_1) D$, is positive.

When populations are fully mixing (i.e. $d_{rs} = v_s$ for all r, s), Metz et al. (1983) derived a simple expression for $r_1(0)$ given by

$$r_1(0) = \mathbb{E} \left[\log \left(\sum_{r=1}^n v_r \lambda_t^r \right) \right] \quad (17)$$

i.e. the temporal log-mean of the spatial arithmetic mean. Owing to the concavity of the log function, Jensens inequality applied to the spatial and temporal averages in (17) yields

$$\log \left(\sum_{r=1}^n v_r \mathbb{E}[\lambda_t^r] \right) > r_1(0) > \sum_{r=1}^n v_r \mathbb{E}[\log \lambda_t^r]. \quad (18)$$

The second inequality implies that dispersal can mediate persistence as $r_1(0)$ can be positive despite all local growth rates $\mathbb{E}[\log \lambda_t^r]$ being negative. Hence, populations can persist even when all patches are sinks, a phenomena that has been observed in the analysis of density-independent models and simulations of density-dependent models (Jansen and Yoshimura 1998; Bascompte et al. 2002; Evans et al. 2013). The first inequality in equation (18), however, implies that dispersal-mediated persistence for well-mixed populations requires that the expected fitness $\mathbb{E}[\lambda_t^r]$ is greater than one in at least one patch.

For partially mixing populations for which $d_{rs} = v_s + \varepsilon \delta_{rs}$, Schreiber (2010) developed first-order approximation of $r_1(0)$ with respect to ε . This approximation coupled with Theorem 5.1 implies that temporal autocorrelations for partially mixing populations can mediate persistence even when the expected fitness $\mathbb{E}[\lambda_t^r]$ is less than one in all patches, a finding related to earlier work by Roy et al. (2005). This dispersal mediated persistence occurs when spatial correlations are sufficiently weak, temporal fluctuations are sufficiently large and positively autocorrelated, and there are sufficiently many patches.

Biological Interpretation 5.4 *Metapopulations with density-dependent growth can stochastically persist despite all local populations being extinction prone in the absence of immigration. Temporal autocorrelations can enhance this effect.*

6 Applications to competing species in space

The roles of spatial and temporal heterogeneity in maintaining diversity is a fundamental problem of practical and theoretical interest in population biology (Chesson 2000a,b; Loreau et al. 2003; Mouquet and Loreau 2003; Davies et al. 2005). To examine the role of both forms of heterogeneity in maintaining diversity of competitive communities, we consider lottery-type models of m competing populations in a landscape consisting of n patches. For these models, competition for vacant space determines the within patch dynamics, while dispersal between the patches couples the local dynamics. After describing a general formulation of these models for an arbitrary number of species with potentially frequency-dependent interactions, we illustrate how to apply

our results to case of two competing species and three competing species exhibiting an intransitive, rock-paper-scissor like dynamic.

6.1 Formulation of the general model

To describe the general model, let X_t^{ir} denote the fraction of patch r occupied by population i at time t . At each time step, a fraction $\varepsilon > 0$ of individuals die in each patch. The sites emptied by the dying individuals get randomly assigned to progeny in the patch. Birth rates within each patch are determined by local pair-wise interactions. Let $\xi_t^{ij}(r)$ be the ‘‘payoff’’ to strategy i interacting with strategy j in patch r at time t . Let

$$\Xi_t(r) = \left(\xi_t^{ij}(r) \right)_{1 \leq i, j \leq m} \tag{19}$$

be the payoff matrix for patch r . The total number of progeny produced by an individual playing strategy i in patch r is $\sum_j \xi_t^{ij} X_t^{jr}$. Progeny disperse between patches with d_{sr} the fraction of progeny dispersing from patch s to patch r . Under these assumptions, the spatial-temporal dynamics of the competing populations are given by

$$X_{t+1}^{ir} = \varepsilon \frac{\sum_s d_{sr} \sum_j \xi_t^{ij}(s) X_t^{js} X_t^{is}}{\sum_s d_{sr} \sum_{j,l} \xi_t^{lj}(s) X_t^{js} X_t^{ls}} + (1 - \varepsilon) X_t^{ir}. \tag{20}$$

Let $A_i(\xi, X)$ be the matrix whose $s - r$ entry is given by

$$\varepsilon \frac{d_{sr} \sum_j \xi^{ij}(s) X^{js}}{\sum_{s'} d_{s'r} \sum_{j,l} \xi^{lj}(s') X^{js'} X^{ls'}}$$

for $r \neq s$, and

$$\varepsilon \frac{d_{sr} \sum_j \xi^{ij}(s) X^{js}}{\sum_{s'} d_{s'r} \sum_{j,l} \xi^{lj}(s') X^{js'} X^{ls'}} + 1 - \varepsilon$$

for $r = s$. With these definitions, (20) takes on the form of our model (1).

To illustrate the insights that can be gained from a persistence analysis of these models, we consider two special cases. The first case is a spatially explicit version of [Chesson and Warner \(1981\)](#)’s lottery model. The second case is a spatial version of a stochastic rock-paper-scissor game. For both of these examples, we assume that a fraction d of all progeny disperse randomly to all patches and the remaining fraction $1 - d$ do not disperse. Under this assumption, we get $d_{sr} = d/(m - 1)$ for $s \neq r$ and $d_{ss} = 1 - d$. These populations are fully mixing when $d = \frac{m-1}{m}$ in which case $d_{sr} = \frac{1}{m}$ for all s, r .

6.2 A spatially-explicit lottery model

The lottery model of [Chesson and Warner \(1981\)](#) assumes that the competing populations do not exhibit frequency dependent interactions. More specifically, the “payoffs” $\xi_t^{ij}(r) = \xi_t^j(r)$ for all i, j are independent of the frequencies of the other species. Consequently, the model takes on a simpler form

$$X_{t+1}^{ir} = \varepsilon \frac{\sum_s d_{sr} \xi_t^i(s) X_t^{is}}{\sum_s d_{sr} \sum_j \xi_t^j(s) X_t^{js}} + (1 - \varepsilon) X_t^{ir} \tag{21}$$

where $d_{sr} = \frac{d}{m-1}$ for $r \neq s$ and $d_{ss} = 1 - d$.

For two competing species (i.e. $m = 2$), the population states $z_1 = (1, \dots, 1, 0, \dots, 0)$ and $z_2 = (0, \dots, 0, 1, \dots, 1)$ correspond to only species 1 and only species 2 occupying the landscape, respectively. The extinction set is $S_0 = \{z_1, z_2\}$. [Theorem 3.1](#) implies that a sufficient condition for stochastic persistence is the existence of positive weights p_1, p_2 such that

$$p_1 r_1(z_1) + p_2 r_2(z_1) > 0 \quad \text{and} \quad p_1 r_1(z_2) + p_2 r_2(z_2) > 0.$$

[Proposition 8.19](#) implies that the long-term growth rate of any invariant measure, with a support bounded away from the extinction set, is equal to zero. In particular, this proposition applies to the subsystems of species 1 and 2, and to the Dirac measures δ_{z_1} and δ_{z_2} , respectively. Therefore $r_1(z_1) = r_2(z_2) = 0$ with probability one. This implies that $r_1(z_1) = r_2(z_2) = 0$. Hence, the persistence criterion simplifies to

$$r_1(z_2) > 0 \quad \text{and} \quad r_2(z_1) > 0.$$

In other words, persistence occurs if each species has a positive invasion rate.

To get some biological intuition from the mutual invasibility criterion, we consider the limiting cases of relatively sedentary populations (i.e. $d \approx 0$) and highly dispersive populations (i.e. $d \approx 1$). In these cases, we get explicit expressions for the realized per-capita growth rates $r_i(z_j)$ that simplify further for short-lived (i.e. $\varepsilon \approx 1$) and long-lived (i.e. $\varepsilon \approx 0$) species. Our analytical results are illustrated numerically in [Fig. 3](#).

6.2.1 Relatively sedentary populations

When populations are completely sedentary (i.e. $d = 0$), the projection matrix $A_2(\xi, z_1)$ corresponding to species 2 trying to invade a landscape monopolized by species 1 reduces to a diagonal matrix whose r -th diagonal entry equals

$$\varepsilon \frac{\xi_r^2(r)}{\xi_r^1(r)} + 1 - \varepsilon.$$

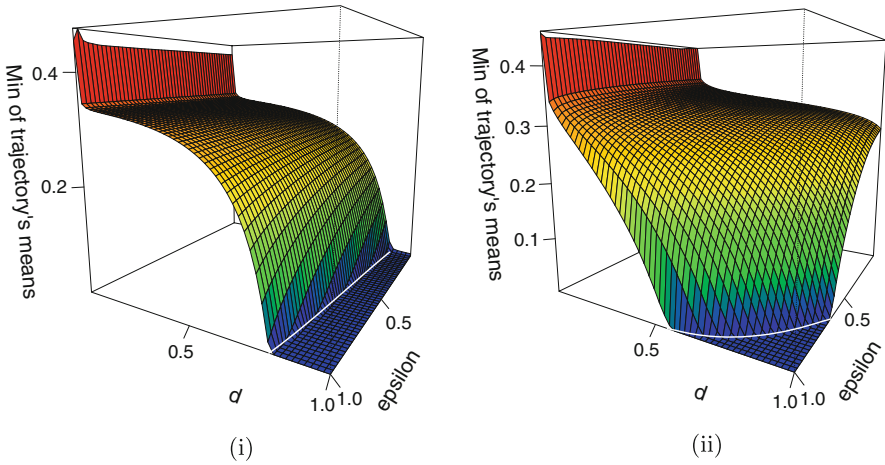


Fig. 3 Effects of dispersal and survival on coexistence of two species. The log-fecundities ξ^i are independent and normally distributed with means $\mu_1 = (5, 0, 5, 0, \dots, 0)$, $\mu_2 = (0, 5, 0, \dots, 0)$ and variances $\sigma_1^2 = \sigma_2^2 = (1, \dots, 1)$ for (I) and $(3, \dots, 3)$ for (II). The white lines correspond to the zero-lines of the respective Lyapunov exponents

The dominant Lyapunov exponent in this limiting case is given by

$$r_2(z_1) = \max_r \mathbb{E} \left[\log \left(\varepsilon \frac{\xi_t^2(r)}{\xi_t^1(r)} + 1 - \varepsilon \right) \right].$$

Proposition 3 from [Benaim and Schreiber \(2009\)](#) implies that $r_2(z_1)$ is a continuous function of d . Consequently, $r_2(z_1)$ is positive for small $d > 0$ provided that $\mathbb{E} \left[\log \left(\varepsilon \frac{\xi_t^2(r)}{\xi_t^1(r)} + 1 - \varepsilon \right) \right]$ is strictly positive for some patch r . Similarly, $r_1(z_2)$ is positive for small $d > 0$ provided that $\mathbb{E} \left[\log \left(\varepsilon \frac{\xi_t^1(r)}{\xi_t^2(r)} + 1 - \varepsilon \right) \right]$ is strictly positive for some patch r . Thus, coexistence for small $d > 0$ occurs if

$$\max_r \mathbb{E} \left[\log \left(\varepsilon \frac{\xi_t^2(r)}{\xi_t^1(r)} + 1 - \varepsilon \right) \right] > 0 \quad \text{and} \quad \max_r \mathbb{E} \left[\log \left(\varepsilon \frac{\xi_t^1(r)}{\xi_t^2(r)} + 1 - \varepsilon \right) \right] > 0.$$

When $\varepsilon \approx 1$ or $\varepsilon \approx 0$, we get more explicit forms of this coexistence condition. When the populations are short-lived ($\varepsilon \approx 1$), the coexistence condition simplifies to $\mathbb{E}[\log \xi_t^1(r)] > \mathbb{E}[\log \xi_t^2(r)]$ and $\mathbb{E}[\log \xi_t^2(s)] > \mathbb{E}[\log \xi_t^1(s)]$ for some patches $r \neq s$. Coexistence requires that each species has at least one patch in which they have a higher geometric mean in their reproductive output.

When the populations are long lived ($\varepsilon \approx 0$) and relatively sedentary ($d \approx 0$), the coexistence condition is

$$\mathbb{E} \left[\frac{\xi_t^2(r)}{\xi_t^1(r)} \right] > 1 \quad \text{and} \quad \mathbb{E} \left[\frac{\xi_t^1(s)}{\xi_t^2(s)} \right] > 1$$

for some patches r, s . Unlike short-lived populations, it is possible that both inequalities are satisfied for the same patch. For example, when the log-fecundities $\log \xi_t^i(r)$ are independent and normally distributed with mean $\mu_i(r)$ and variance $\sigma_i^2(r)$, the coexistence conditions is

$$\mu_2(r) - \mu_1(r) + \frac{\sigma_1^2(r) + \sigma_2^2(r)}{2} > 1$$

for some patch r and

$$\mu_1(s) - \mu_2(s) + \frac{\sigma_1^2(s) + \sigma_2^2(s)}{2} > 1$$

for some patch s . Both conditions can be satisfied in the same patch r provided that $\sigma_1(r)$ or $\sigma_2(r)$ is sufficiently large.

Biological Interpretation 6.1 *For relatively sedentary populations, coexistence occurs if each species has a patch it can invade when rare. If the populations are also short-lived, coexistence requires that each species has a patch in which it is competitively dominant. Alternatively, if populations are also long-lived, regional coexistence may occur if species coexist locally within a patch due to the storage effect. For uncorrelated and log-normally distributed fecundities, this within-patch storage effect occurs if the difference in the mean log-fecundities is sufficiently smaller than the net variance in the log-fecundities.*

6.2.2 Well-mixed populations

For populations that are highly dispersive (i.e. $d = \frac{m-1}{m}$), the spatially explicit Lottery model reduces to a spatially implicit model where

$$r_1(z_2) = \mathbb{E} \left[\log \left(\varepsilon \frac{\sum_r \xi_t^1(r)}{\sum_r \xi_t^2(r)} + 1 - \varepsilon \right) \right] \text{ and}$$

$$r_2(z_1) = \mathbb{E} \left[\log \left(\varepsilon \frac{\sum_r \xi_t^2(r)}{\sum_r \xi_t^1(r)} + 1 - \varepsilon \right) \right].$$

For short lived populations ($\varepsilon = 1$), these long-term growth rates simplify to

$$r_1(z_2) = \mathbb{E} \left[\log \sum_r \xi_t^1(r) \right] - \mathbb{E} \left[\log \sum_r \xi_t^2(r) \right] \text{ and}$$

$$r_2(z_1) = \mathbb{E} \left[\log \sum_r \xi_t^2(r) \right] - \mathbb{E} \left[\log \sum_r \xi_t^1(r) \right].$$

Since $r_1(z_2) = -r_2(z_1)$, the persistence criterion that $r_1(z_2) > 0$ and $r_2(z_1) > 0$ is not satisfied generically.

Alternatively, for long-lived populations ($\varepsilon \approx 0$), the invasion rates of well-mixed populations becomes to first-order in $\varepsilon > 0$:

$$r_1(z_2) \approx \varepsilon \left(\mathbb{E} \left[\frac{\sum_r \xi_r^1(r)}{\sum_r \xi_r^2(r)} \right] - 1 \right) \text{ and}$$

$$r_2(z_1) \approx \varepsilon \left(\mathbb{E} \left[\frac{\sum_r \xi_r^2(r)}{\sum_r \xi_r^1(r)} \right] - 1 \right).$$

We conjecture that this coexistence condition is less likely to be met than the coexistence condition for relatively sedentary populations. To see why, consider a small variance approximation of these invasion rates. Assume that $\xi_r^i = \bar{\xi}^i + \eta Z_r^i(r)$ where $Z_r^i(r)$ are independent and identically distributed in i, r and $\mathbb{E}[Z_r^i(r)] = 0$ for all i, r . Let $\sigma^2 = \mathbb{E}[(Z_r^i(r))^2]$. A second order Taylor's approximation in η yields the following approximation of the (rescaled) long-term growth rates for well-mixed populations

$$\mathbb{E} \left[\frac{\sum_r \xi_r^1(r)}{\sum_r \xi_r^2(r)} \right] - 1 \approx \frac{\bar{\xi}^1}{\bar{\xi}^2} + \frac{\bar{\xi}^1 \sigma^2 / n}{(\bar{\xi}^2)^3} - 1 \quad (22)$$

and the following approximation for relatively sedentary populations

$$\max_r \mathbb{E} \left[\frac{\xi_r^1(r)}{\xi_r^2(r)} \right] - 1 \approx \frac{\bar{\xi}^1}{\bar{\xi}^2} + \frac{\bar{\xi}^1 \sigma^2}{(\bar{\xi}^2)^3} - 1. \quad (23)$$

Since (23) is greater than (22), persistence is more likely for relatively sedentary populations in this small noise limit.

Biological Interpretation 6.2 *Short-lived and highly dispersive competitors do not satisfy the coexistence condition. Long-lived and highly-dispersive competitors may coexist. However, coexistence appears to be less likely than for sedentary populations as spatial averaging reduces the temporal variability experienced by both populations and, thereby, weakens the storage effect.*

6.3 The rock-paper-scissor game

In the last few years the rock-paper-scissor game, which might initially seem to be of purely theoretical interest, has emerged as playing an important role in describing the behavior of various real-world systems. These include the evolution of alternative male mating strategies in the side-blotched lizard *Uta Stansburiana* (Sinervo and Lively 1996), the *in vitro* evolution of bacterial populations (Kerr et al. 2002; Nahum et al. 2011), the *in vivo* evolution of bacterial populations in mice (Kirkup and Riley 2004), and the competition between genotypes and species in plant communities (Lankau and Strauss 2007; Cameron et al. 2009). More generally, the rock-scissors-paper game – which is characterized by three strategies R, P and S, which satisfy the non-transitive relations: P beats R (in the absence of S), S beats P (in the absence of R), and R beats S (in the absence of P) – serves as a simple prototype for studying the dynamics

of more complicated non-transitive systems (Buss and Jackson 1979; Paquin and Adams 1983; May and Leonard 1975; Schreiber 1997; Schreiber and Rittenhouse 2004; Vandermeer and Pascual 2005; Allesina and Levine 2011). Here, we examine a simple spatial version of this evolutionary game in a fluctuating environment.

Let $x_t^1(r)$, $x_t^2(r)$, and $x_t^3(r)$ be the frequencies of the rock, paper, and scissor strategies in patch r , respectively. All strategies in patch r receive a basal payoff of a_t^r at time t . Winners in an interaction in patch r receive a payoff of b_t^r while losers pay a cost c_t^r . Thus, the payoff matrix (19) for the interacting populations in patch r is

$$\Xi_t(r) = a_t^r + \begin{pmatrix} 0 & -c_t^r & b_t^r \\ b_t^r & 0 & -c_t^r \\ -c_t^r & b_t^r & 0 \end{pmatrix}.$$

We continue to assume that the fraction of progeny dispersing from patch s to patch r equals $d/(m - 1)$ for $s \neq r$ and $1 - d$ otherwise.

Our first result about the rock-paper-scissor model is that it exhibits a heteroclinic cycle in S_0 between the three equilibria $E_1 = (1, \dots, 1, 0, \dots, 0, 0, \dots, 0)$, $E_2 = (0, \dots, 0, 1, \dots, 1, 0, \dots, 0)$ and $E_3 = (0, \dots, 0, 0, \dots, 0, 1, \dots, 1)$. For two vectors $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, we write $x > y$ if $x_i \geq y_i$ for all i with at least one strict inequality.

Proposition 6.3 *Assume $d, \varepsilon \in (0, 1]$ and $a_t^r > c_t^r$, $\log a_t^r, \log c_t^r, \log b_t^r \in [-M, M]$ with probability one for some $M > 0$. If $x_0^1 > (0, \dots, 0)$ and $x_0^2 > (0, \dots, 0)$ and $x_0^3 = (0, \dots, 0)$, then $\lim_{t \rightarrow \infty} x_t = E_2$ with probability one. If $x_0^1 > (0, \dots, 0)$ and $x_0^3 > (0, \dots, 0)$ and $x_0^2 = (0, \dots, 0)$, then $\lim_{t \rightarrow \infty} x_t = E_1$ with probability one. If $x_0^2 > (0, \dots, 0)$ and $x_0^3 > (0, \dots, 0)$ and $x_0^1 = (0, \dots, 0)$, then $\lim_{t \rightarrow \infty} x_t = E_3$ with probability one.*

Proof It suffices to prove the assertion for the case in which $x_0^1 > (0, \dots, 0)$ and $x_0^2 > (0, \dots, 0)$ and $x_0^3 = (0, \dots, 0)$. Let $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$. Our assumptions $b_t^r > 0$ and $a_t^r > c_t^r > 0$ imply there exists $\eta > 0$ such that $A_2(\xi_{t+1}, X_t) \gg \exp(\eta)A_1(\xi_{t+1}, X_t)$ with probability one. It follows that

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{1}{t} \log \|X_t^1\| &= \limsup_{t \rightarrow \infty} \frac{1}{t} \log \|X_0^1 A_1(\xi_1, X_0) \dots A_1(\xi_t, X_{t-1})\| \\ &\leq \limsup_{t \rightarrow \infty} \frac{1}{t} \log \|X_0^1 A_2(\xi_1, X_0) \dots A_2(\xi_t, X_{t-1})\| - \eta \\ &= \limsup_{t \rightarrow \infty} \frac{1}{t} \log \|\mathbf{1} A_2(\xi_1, X_0) \dots A_2(\xi_t, X_{t-1})\| - \eta \\ &\leq -\eta \end{aligned}$$

where the last two lines follow from Proposition 8.16 and its Corollary 8.17. Hence, $\lim_{t \rightarrow \infty} \|X_t^1\| = 0$ as claimed. □

Proposition 6.3 implies that for any $x \in S_0$ and $1 \leq i \leq 3$, $r_i(x) = r_i(E_j)$ for some $1 \leq j \leq 3$. Hence, the persistence criterion of Theorem 3.1 requires $p_1, p_2, p_3 > 0$ such that

$$\sum_i p_i r_i(E_j) > 0 \quad \text{for all } 1 \leq j \leq 3.$$

A standard algebraic calculation shows that this persistence criterion is satisfied if and only if

$$r_2(E_1)r_3(E_2)r_1(E_3) > -r_3(E_1)r_1(E_2)r_2(E_3)$$

i.e. the product of the positive invasion rates is greater than the absolute value of the product of the negative invasion rates. The symmetry of our model implies that all the positive invasion rates are equal and all the negative invasion rates are equal. Hence, coexistence requires

$$r_2(E_1) > -r_3(E_1).$$

As for the case of two competing species, we can derive more explicit coexistence criteria when the populations are relatively sedentary (i.e. $d \approx 0$) or the populations are well-mixed (i.e. $d = \frac{m}{m-1}$). For relatively sedentary populations, coexistence requires

$$\max_r \mathbb{E} \left[\log \left(1 - \varepsilon + \varepsilon \frac{a_t^r + b_t^r}{a_t^r} \right) \right] > -\max_r \mathbb{E} \left[\log \left(1 - \varepsilon + \varepsilon \frac{a_t^r - c_t^r}{a_t^r} \right) \right].$$

For long-lived populations, this coexistence criterion simplifies further to

$$\max_r \mathbb{E} \left[\frac{b_t^r}{a_t^r} \right] > \min_r \mathbb{E} \left[\frac{c_t^r}{a_t^r} \right].$$

Alternatively, when the populations are well-mixed, coexistence requires

$$\mathbb{E} \left[\log \left(1 - \varepsilon + \varepsilon \frac{\sum_r a_t^r + b_t^r}{\sum_r a_t^r} \right) \right] > -\max_r \mathbb{E} \left[\log \left(1 - \varepsilon + \varepsilon \frac{\sum_r a_t^r - c_t^r}{\sum_r a_t^r} \right) \right].$$

For long-lived populations, this coexistence criterion simplifies further to

$$\mathbb{E} \left[\frac{\sum_r b_t^r}{\sum_r a_t^r} \right] > \min_r \mathbb{E} \left[\frac{\sum_r c_t^r}{\sum_r a_t^r} \right].$$

Biological Interpretation 6.4 *For relatively sedentary populations, coexistence only requires that average benefits (relative to the base payoff) in one patch is greater than the average costs (relative to the base payoff) in another patch. Negative correlations between benefits b_t^r and basal payoffs a_t^r promote coexistence. For highly dispersive species whose base payoffs are constant in space in time (i.e. $a_t^r = a$ for all t, r), coexistence requires the spatially and temporally averaged benefits of interactions exceed the spatially and temporally averaged costs of interactions.*

7 Discussion

Understanding the conditions that ensure the long-term persistence of interacting populations is of fundamental theoretical and practical importance in population biology. For deterministic models, coexistence naturally corresponds to an attractor bounded away from extinction. Since populations often experience large perturbation, many authors have argued that the existence of a global attractor (i.e. permanence or uniform persistence) may be necessary for long-term persistence (Hofbauer and Sigmund 1998; Smith and Thieme 2011). Most populations experience stochastic fluctuations in their demographic parameters (May 1973) which raises the question (May 1973, p. 621) “How are the various usages of the term [persistence] in deterministic and stochastic circumstances related?” Only recently has it been shown that the deterministic criteria for permanence extend naturally to criteria for stochastic persistence in stochastic difference and differential equations (Benaïm et al. 2008; Schreiber et al. 2011). These criteria assume that the populations are unstructured (i.e. no differences among individuals) and environmental fluctuations are temporally uncorrelated. However, many populations are structured as highlighted in a recent special issue in *Theoretical Population Biology* (Tuljapurkar et al. 2012) devoted to this topic. Moreover, many environmental factors such as temperature and precipitation exhibit temporal autocorrelations (Vasseur and Yodzis 2004). Here, we prove that by using long-term growth rates when rare, the standard criteria for persistence extend to models of interacting populations experiencing correlated as well as uncorrelated environmental stochasticity, exhibiting within population structure, and any form of density-dependent feedbacks. To illustrate the utility of these criteria, we applied them to persistence of predator-prey interactions in auto-correlated environments, structured populations with overcompensating density-dependence, and competitors in spatially structured environments.

Mandelbrot (1982) proposed that environmental signals commonly found in nature may be composed of frequencies f that scale according to an inverse power law $1/f^\beta$. With this scaling, uncorrelated (i.e. white) noise corresponds to $\beta = 0$, positively auto-correlated (i.e. red or brown) noise corresponds to $\beta > 0$, and negatively auto-correlated (e.g. blue) noise corresponds to $\beta < 0$. Many environmental signals important to ecological processes including precipitation, mean air temperature, degree days, and seasonal indices exhibit positive β exponents (Vasseur and Yodzis 2004). Consistent with prior work on models with compensating density dependence (Roughgarden 1975; Johst and Wissel 1997; Petchey 2000), we found that positive autocorrelations in the maximal per-capita growth rate of species increases the long-term variability in their densities. If this species is the prey for a predatory species, we showed that this increased variability in prey densities reduced a predator’s realized per-capita growth rate when rare. Hence, positive autocorrelations may impede predator-prey coexistence. In contrast, negative autocorrelations, possibly due to a biotic feedback between the prey species and its resources, may facilitate coexistence by reducing variation in prey densities and, thereby, increase the predator’s growth rate when rate. These results are qualitatively consistent with prior results that positive-autocorrelations in predator-prey systems can increase variation in prey and predator densities when they coexist (Collie and Spencer 1994; Ripa and Ives 2003). Specifically, in a simulation

study of predator-prey interactions in pelagic fish stocks, [Collie and Spencer \(1994\)](#) found reddened noise resulted in predator-prey densities “to shift between high and low equilibrium levels” and, thereby, increase variability in their abundances. Similarly, using linear approximations, [Ripa and Ives \(2003\)](#) found that environmental autocorrelations increased the amplitude of populations cycles. All of these results, however, stem from the per-capita growth rate of the predator being an increasing, concave function of prey density. Changes in concavity (e.g. a type-III functional response) could produce an opposing result: increased variability in prey densities may facilitate predator invasions. A more detailed analysis of this alternative is still needed.

Classical stochastic demography theory ([Tuljapurkar 1990](#); [Boyce et al. 2006](#)) considers population growth rates in the absence of density-dependent feedbacks. Our results for populations experiencing negative-density dependence show that stochastic persistence depends on the population’s long-term growth rate $r(0)$ when rare. Hence, applying stochastic demography theory to $r(0)$ provides insights into how environmental stochasticity interacts with population structure to determine stochastic persistence. For example, a fundamental result from stochastic demography is that positive, within-year correlations between vital rates decreases $r(0)$ and thereby may thwart stochastic persistence, a result consistent with our analysis of the stochastic LPA model for flour beetle dynamics. Stochastic demography theory also highlights that temporal autocorrelations can have subtle effects on $r(0)$. In particular, for a density-independent version of the metapopulation model considered here, [Schreiber \(2010\)](#) demonstrated that positive temporal autocorrelations can increase the metapopulation growth rate $r(0)$ when rare for partially mixing populations, a prediction consistent with laboratory experiments ([Matthews and Gonzalez 2007](#)) and earlier theoretical work ([Roy et al. 2005](#)). In contrast, [Tuljapurkar and Haridas \(2006\)](#) found that negative temporal autocorrelations between years with and without fires increased the realized per-capita growth rate $r(0)$ for models of the endangered herbaceous perennial *Lomatium bradshawii*. Our results imply that these results also apply to models accounting for density-dependence.

Spatial heterogeneity of populations has been shown theoretically and empirically to have an effect on coexistence of competitive species (see e.g. [Amarasekare \(2003\)](#) or [Chesson \(2000b\)](#) for a review). Coexistence requires species to exhibit niche differentiation that decrease the interspecific competition ([Chesson 2000a](#)). In a fluctuating environment, these niches can arise as differential responses to temporal variation ([McGehee and Armstrong 1977](#); [Armstrong and McGehee 1980](#); [Chesson 2000a,b](#)), spatial variation ([May and Hassell 1981](#); [Chesson 2000a,b](#); [Snyder and Chesson 2003](#)), or a combination of both forms of variation ([Chesson 1985](#); [Snyder 2007, 2008](#)). For the spatial lottery model where species disperse between a finite number of patches and compete for micro sites within these patches, our coexistence criterion applies, and reduces to the *mutual invasibility* criterion. Although [Chesson \(1985\)](#) proved this result in the limit of an infinite number of patches with temporally uncorrelated fluctuations, our result is less restrictive as the number of patches can be small and temporal fluctuations can be autocorrelated. Using this mutual invasibility criterion, we derive explicit coexistence criteria for relatively sedentary populations and highly dispersive populations. In the former case, coexistence occurs if each species has a patch

it can invade when rare. For short-lived populations, coexistence requires that each species has a patch in which it is competitively dominant. Alternatively, for long-lived populations, regional coexistence may occur if species coexist locally within a patch due to the storage effect (Chesson and Warner 1981; Chesson 1982, 1994) in the *one patch* case. For highly dispersive populations, the coexistence criterion is only satisfied if populations exhibit overlapping generations, a conclusion consistent with (Chesson 1985). By providing the first mathematical confirmation of the mutual inviability criterion for the spatial lottery model with spatial and temporal variation, our result opens the door for mathematically more rigorous investigations in understanding the relative roles of temporal variation, spatial heterogeneity, and dispersal on coexistence.

For lottery models with three or more species, persistence criteria are more subtle and inviability of all sub communities isn't always sufficient (May and Leonard 1975). For example, in rock-paper-scissor communities where species 2 displaces species 1, 3 displaces 2, and 1 displaces 3, all sub communities which consist of a single species are inviable by another, but coexistence may not occur (Hofbauer and Sigmund 1998; Schreiber and Killingback 2013). For the deterministic models, coexistence requires that the geometric mean of the benefits of pair-wise interactions exceeds the costs of these interactions (Schreiber and Killingback 2013). Schreiber et al. (2011) and Schreiber and Killingback (2013) studied these interactions in models *separately* accounting for temporal fluctuations or spatial heterogeneity. In both cases, temporal heterogeneity or spatial heterogeneity can individually promote coexistence. Here we extend these result to intransitive communities experiencing both spatial heterogeneity and temporal fluctuations, thereby unifying this prior work. Our persistence criterion reduces to: the geometric mean of the positive long-term, low-density growth rates of each species (e.g. invasion rate of rock to scissor) is greater than the geometric mean of the absolute values of the negative, long-term, low-density growth rates (e.g. invasion rate of rock to paper). For relatively sedentary populations, coexistence only requires that average benefits (relative to the base payoff) in one patch is greater than the average costs in another patch. Moreover, negative correlations between benefits and basal payoffs promote coexistence. For highly dispersive species, coexistence requires the spatially and temporally averaged benefits of interactions exceed the spatially and temporally averaged costs of interactions, assuming that base payoffs are constant in space and time.

The theory of stochastic population dynamics is confronted with many, exciting challenges. First, our persistence criterion requires every sub community (as represented by an ergodic invariant measure supporting a subset of species) is inviable by at least one missing species. While this inviability condition in general isn't sufficient for coexistence, understanding when it is sufficient remains a challenging open question. For example, it should be sufficient for most food chain models (see the argument for deterministic models in Schreiber 2000), non-interacting prey species sharing a common predator (see the argument for deterministic models in Schreiber 2004), and species competing for a single resource species. However, finding a simple criterion underlying these examples is lacking. Second, while we have provided a sufficient condition for stochastic persistence, it is equally important to develop sufficient conditions for the asymptotic exclusion of one or more species with positive

probability. In light of the deterministic theory, a natural conjecture in this direction is the following: if there exist non-negative weights p_1, \dots, p_k such that

$$\sum_i p_i r_i(x) < 0$$

for every population state x in the extinction set S_0 , then there exist positive initial conditions such that X_t asymptotically approaches S_0 with positive probability. [Benaïm et al. \(2008\)](#) proved a stronger version of this conjecture for stochastic differential equation models where the diffusion term is small and the populations are unstructured. However, it is not clear whether these methods carry over to models with “large” noise or population structure. Another important challenge is relaxing the compactness assumption **H4** for our stochastic persistence results. While this assumption is biologically realistic (i.e. populations always have an upper limit on their size), it is theoretically inconvenient as many natural models of environmental noise have non-compact distributions (e.g. log-normal or gamma distributions). One promising approach developed by [Benaïm and Schreiber \(2009\)](#) for structured models of single species is identifying Lyapunov-like functions that decrease on average when population densities get large. Finding sufficient conditions for “stochastic boundedness” is only half of the challenge, extending the stochastic persistent results to these “stochastically bounded” models will require additional innovations. Finally, and most importantly, there is a desperate need to develop more tools to analytically approximate or directly compute the long-term growth rates $r_i(\mu)$ when rare. One promising approach is [Pollicott \(2010\)](#)’s recently derived power series representation of Lyapunov exponents.

8 Proof of Theorems 3.1 and 3.4

This Section proves Theorem 3.4 from which Theorem 3.1 follows. Sections 8.1 and 8.2 lead to the statement of Theorem 8.11 which is equivalent to Theorem 3.4. The rest of the section is dedicated to the proof of Theorem 8.11. More specifically, in Sect. 8.1, we recast our stochastic model (1) and our main hypothesis in Arnold’s framework of random dynamical system ([Arnold 1998](#); [Bhattacharya and Majumdar 2007](#)). The purpose of this recasting is to write explicitly the underlying dynamics of the matrix products (3) in order to use the Random Perron-Frobenius Theorem ([Ruelle \(1979a\)](#)), a key element in the proof of Theorem 3.4. The Random Perron-Frobenius Theorem requires this underlying dynamics to be invertible which is, a priori, not the case here. Therefore, in Sect. 8.2, we extend the underlying dynamics to an invertible dynamic on the trajectory space and state Theorem 8.11 which is equivalent to Theorem 3.4. Working in the Arnold’s framework and extending the dynamic to the trajectory space requires three forms of notation (i.e. main text, random dynamical system and trajectory space) that are summarized in Table 1. In Sect. 8.3, we prove basic results about the average per-capita growth rates r_i . In Sect. 8.4, we prove several basic results about occupational measures and their weak* limit points. These

Table 1 Notation for the probabilistic, RDS, and trajectory space formulations of the population dynamics

Probabilistic formulation		RDS formulation		Trajectory space formulation
Environmental state space E , a Polish space		$\Omega = E^{\mathbb{Z}}$, the space of all sequences of environment state.		Ω
$\xi_t \in E$, the environment state at time t		$\omega = (\dots, \omega_{-1}, \omega_0, \omega_1, \dots) \in \Omega$ is a sequence of environment states, i.e. $e_t = \xi_t$		$\omega = (\dots, \omega_{-1}, \omega_0, \omega_1, \dots)$
State space \mathbb{R}_+^n	\leftarrow	$\Omega \times \mathbb{R}_+^n$	\leftarrow	$\Gamma_+ \subset \Omega \times (\mathbb{R}_+^n)^{\mathbb{Z}}$
$x_0 = p_2(\omega, x_0)$	p_2	$(\omega, x_0) = \pi_0(\gamma)$	π_0	$\gamma = (\omega, \{x_t\}_{t \in \mathbb{Z}})$
Dynamics For one time step: $X_1 = x_0 A(\xi_1, x_0)$		$\Phi(\omega, x_0) = (\theta(\omega), x_0 A(\omega_0, x_0))$		Θ , the shift operator on Γ_+
For t time steps: $X_t = x_0 A(\xi_1, x_0) \cdots A(\xi_t, X_{t-1})$	"="	$p_2(\Phi^t(\omega, x_0))$	$=$	$p_2(\pi_0(\Theta^t(\gamma)))$
Empirical measures Π_t^x For a Borel set $B \in \mathbb{R}_+^n$: $\Pi_t^{x_0}(B)$	"="	$\Lambda_t(\omega, x_0)(\Omega, B)$	$=$	$\tilde{\Lambda}_t(\gamma)(\Omega, B)$
Invariant measures $\text{Inv} := \{ \mu \text{ satisfying Def. 3.2} \}$	\leftarrow	$\text{Inv}_{\mathbb{Q}}(\Phi)$	\leftarrow	$\text{Inv}_{\mathbb{Q}}(\Theta)$
$h(\mu)$	h	$\mu = \pi_0^*(\tilde{\mu})$	π_0^*	$\tilde{\mu}$
$h^{-1}(\text{Inv})$	$=$	$\text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times \bar{V})$	\supset	$\pi_0^*(\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+))$
Long-term growth rates $r_i(x_0)$ For $\mu \in \text{Inv}$: $r_i(h(\mu))$	"=" $=$	$r_i(\omega, x_0)$ For $\mu \in \text{Inv}_{\mathbb{Q}}(\Phi)$: $r_i(\mu)$	$=$ $=$	$r_i(\gamma)$ For $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)$: $r_i(\tilde{\mu})$

"=" denotes equivalence when ω is randomly drawn from Ω

basic results are proven for the extended state space. Proposition 8.10 and Lemma 8.20 translate these results to non-extended state space. A proof of Theorem 8.11 is provided in Sect. 8.5.

8.1 Random dynamical systems framework

To prove our main result, it is useful to embed (1) and assumptions **H1-H4** within Arnold’s general framework of random dynamical systems. Let $\Omega = E^{\mathbb{Z}}$ be the set of possible environmental trajectories, $\mathcal{F} = \mathcal{E}^{\mathbb{Z}}$ be the product σ -algebra on Ω ,

$\theta : \Omega \mapsto \Omega$ be the shift operator defined by $\theta(\{\omega_t\}_{t \in \mathbb{Z}}) = \{\omega_{t+1}\}_{t \in \mathbb{Z}}$, and \mathbb{Q} be the probability measure on Ω satisfying

$$\mathbb{Q}(\{\omega \in \Omega : \omega_t \in E_0, \dots, \omega_{t+k} \in E_k\}) = \mathbb{P}(\xi_0 \in E_0, \dots, \xi_k \in E_k)$$

for any Borel sets $E_0, \dots, E_k \subset E$. Since E is a Polish space, the space Ω endowed with the product topology is Polish as well. Therefore, by the Kolmogorov consistency theorem, the probability measure \mathbb{Q} is well defined, and by a theorem of Rokhlin (1964), θ is ergodic with respect to \mathbb{Q} . Randomness enters by choosing randomly a point $\omega = \{\omega_t\}_{t \in \mathbb{Z}} \in \Omega$ with respect to the probability distribution \mathbb{Q} and defining the environmental state at time t as ω_t .

In this framework, the dynamics (1) takes on the form

$$\begin{cases} X_{t+1}(\omega, x) = X_t(\omega, x)A(\omega_t, X_t(\omega, x)) \\ X_0(\omega, x) = x \in S. \end{cases} \tag{24}$$

We call (24), the *random dynamical system determined by* (θ, \mathbb{P}, A) .

Define the skew product

$$\begin{aligned} \Phi : \Omega \times \mathbb{R}_+^n &\rightarrow \Omega \times \mathbb{R}_+^n \\ (\omega, x) &\mapsto (\theta(\omega), xA(\omega_0, x)) \end{aligned}$$

associated with the dynamics (24) and define the projection maps $p_1 : \Omega \times \mathbb{R}^n \rightarrow \Omega$ and $p_2 : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ by $p_1(x, \omega) = \omega$ and $p_2(x, \omega) = x$. Let Φ^t denote the composition of Φ with itself t times, for $t \in \mathbb{N}$. Remark 1.1.8 in Arnold (1998) implies that the random dynamical system (24) is characterized by the skew product Φ and vice versa. In particular, note that $X_{t+1}(\omega, x) = p_2 \circ \Phi^{t+1}(\omega, x)$ for $x \in S$ and $\omega \in \Omega$. Working with Φ allows the use of the discrete dynamical system theory.

Definition 8.1 A compact set $K \subset \Omega \times \mathbb{R}_+^n$ is a *global attractor for* Φ if there exists a neighborhood V of K such that

- (i) for all $(\omega, x) \in \Omega \times \mathbb{R}_+^n$, there exist $T \in \mathbb{N}$ such that $\Phi^t(\omega, x) \in V$ for all $t \geq T$;
- (ii) $\Phi(V) \subset V$ and $K = \bigcap_{t \in \mathbb{N}} \Phi^t(V)$.

In this random dynamical systems framework, our assumptions **H1** and **H4** take on the form

H1': Ω is a compact space, \mathbb{Q} is a Borel probability measure, and θ is an invertible map that is ergodic with respect to \mathbb{Q} , i.e. for all Borel set $B \subset \Omega$, such that $\theta^{-1}(B) = B$, we have $\mathbb{Q}(B) \in \{0, 1\}$.

H4': There exists a global attractor $K \subset \Omega \times \mathbb{R}_+^n$ for Φ .

Assumptions **H2–H3** do not need to be rewritten in the new framework. Since every ergodic stationary processes on a Polish space can be described as an ergodic measure preserving transformation (Kolmogorov consistency theorem and Rokhlin theorem), assumption **H1'** is less restrictive than **H1**. Assumption **H4'** is simply restatement of assumption **H4** in the random dynamical systems framework.

To state Theorem 3.4 in this random dynamical systems framework, we define invariant measures for the random dynamical system (24). We follow the definition given by Arnold (1998). First, recall some useful definitions and notations. Let M be a metric space, and let $\mathcal{P}(M)$ be the space of Borel probability measures on M endowed with the weak* topology. If M' is also a metric space and $f : M \rightarrow M'$ is Borel measurable, then the induced linear map $f^* : \mathcal{P}(M) \rightarrow \mathcal{P}(M')$ associates with $\nu \in \mathcal{P}(M)$ the measure $f^*(\nu) \in \mathcal{P}(M')$ defined by

$$f^*(\nu)(B) = \nu(f^{-1}(B))$$

for all Borel sets B in M' . If $\theta : M \rightarrow M$ is a continuous map, a measure $\nu \in \mathcal{P}(M)$ is called θ -invariant if $\nu(\theta^{-1}(B)) = \nu(B)$ for all Borel sets $B \in M$. A set $B \subset M$ is positively invariant if $\theta(B) \subset B$. For every positively invariant compact set B , let $\text{Inv}(\theta)(B)$ be the set of all θ -invariant measures supported on B .

Definition 8.2 A probability measure μ on $\Omega \times \mathbb{R}_+^n$ is invariant for the random dynamical system (24) if

- (i) $\mu \in \text{Inv}(\Phi)(\Omega \times \mathbb{R}_+^n)$,
- (ii) $p_1^*(\mu) = \mathbb{Q}$, i.e. for all Borel sets $D \subset \Omega$, $\mu(D \times \mathbb{R}_+^n) = \mathbb{Q}(D)$.

For any positively invariant set $\Omega \times C$ where $C \subset \mathbb{R}_+^n$ is compact, $\text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times C)$ is the set of all measures μ satisfying (i) and (ii) such that $\mu(\Omega \times C) = 1$.

In words, a probability measure μ is invariant for the random dynamical system (24) if it is invariant for the skew product Φ and if its first marginal is the probability \mathbb{Q} on Ω .

The following result is a consequence of Theorem 1.5.10 in Arnold (1998). In fact, the topology defined in his definition 1.5.3 is finer than the weak* topology on the set of all probability measures on $\Omega \times C$.

Proposition 8.3 If $C \subset \mathbb{R}_+^n$ is a positively invariant compact set, then $\text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times C)$ is a nonempty, convex, compact subset of $\mathcal{P}(\Omega \times \mathbb{R}_+^n)$.

The main assumption in Theorem 3.4 deals with the long-term growth rates which characterize, in some sense, the long-term behavior of random matrix products (see Definition 3.3). In order to define those products in the new framework, let $\mathbf{M}_d(\mathbb{R})$ be the set of all $d \times d$ matrices over \mathbb{R} and consider the maps $A_i : \Omega \times S \rightarrow \mathbf{M}_{n_i}(\mathbb{R})$, defined by

$$A_i(\omega, x) = A_i(\omega_0, x).$$

While our choice of notation here differs slightly from the main text, this choice simplifies the proof. We write

$$A_i^t(\omega, x) := A_i(\omega, x)A_i(\Phi(\omega, x)) \cdots A_i(\Phi^{t-1}(\omega, x)), \tag{25}$$

with the convention that $A_i^0(\omega, x) = \text{id}$, the identity matrix.

Then, for each $i \in \{1, \dots, m\}$, the asymptotic growth rate of the product (25) associated with $(\omega, x) \in \Omega \times \mathbb{R}_+^n$ is

$$r_i(\omega, x) := \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|A_i^t(\omega, x)\|,$$

which is finite, due to assumptions **H3** and **H4'**. According to Definition 8.2, the invasion rate of species i with respect to an invariant measure $\mu \in \text{Inv}_{\mathbb{Q}}(\Phi)$ is

$$r_i(\mu) := \int_{\Omega \times \mathbb{R}_+^n} r_i(\omega, x) \mu(d\omega, dx).$$

Remark 8.4 Note that for any $x \in \mathbb{R}_+^n$, the random variable $r_i(x)$ defined by (4) is equal in distribution to the random variable $r_i(\cdot, x)$. Also by definition of \mathbb{Q} and Φ there is a bijection, say h , between the set $\text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times \mathbb{R}_+^n)$ and the set of measures defined in Definition 3.2. Moreover the invasion rate with respect to an invariant measure is invariant by h , i.e. for all $\mu \in \text{Inv}_{\mathbb{Q}}(\Phi)$, $r_i(\mu) = r_i(h(\mu))$.

Given a point $(\omega, x) \in \Omega \times \mathbb{R}_+^n$, let $\Pi_t(\omega, x)$ denote the empirical occupation measure of the trajectory $\{X_s(\omega, x)\}_{s \geq 0}$ at time t defined by

$$\Pi_t(\omega, x) := \frac{1}{t} \sum_{s=0}^{t-1} \delta_{X_s(\omega, x)}.$$

For each Borel set $B \subset \mathbb{R}_+^n$, the random variable $\Pi_t^x(B)$ given by (2) is equal in distribution to the random variable $\Pi_t(\cdot, x)(B)$.

For all $\eta > 0$, recall that $S_\eta := \{x \in \mathbb{R}_+^n : \|x^i\| \leq \eta \text{ for some } i\}$. We can now rephrase Theorem 3.4 in the framework of random dynamical systems.

Theorem 8.5 *If one of the following equivalent conditions hold*

- (i) $r_*(\mu) := \max_{0 \leq i \leq m} r_i(\mu) > 0$ for every probability measure $\mu \in \text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times S_0)$, or
- (ii) there exist positive constants p_1, \dots, p_m such that

$$\sum_i p_i r_i(\mu) > 0$$

- for every ergodic probability measure $\mu \in \text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times S_0)$,
- (iii) there exist positive constants p_1, \dots, p_m such that

$$\sum_i p_i r_i(\omega, x) > 0$$

for every $x \in S_0$ and \mathbb{Q} -almost all $\omega \in \Omega$,

then for all $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\limsup_{t \rightarrow \infty} \Pi_t(\omega, x)(S_\eta) \leq \varepsilon \text{ for } \mathbb{Q}\text{-almost all } \omega,$$

whenever $x \in \mathbb{R}_+^n \setminus S_0$.

Remark 8.4 implies that Theorem 8.5 and Theorem 3.4 are equivalent. The remainder of the section is devoted to prove Theorem 8.5.

8.2 Trajectory space

The key element of the proof of Theorem 8.5 is Proposition 8.13 due to Ruelle (1979a) in which it is crucial that the map Φ is an homeomorphism. However, the map Φ is, a priori, not invertible. To circumvent this issue, we extend the dynamics induced by Φ to an invertible map on the space of possible trajectories. Then, we state an equivalent version of Theorem 8.5 in this larger space that we prove in Sect. 8.5.

By definition of the global attractor K , there exist a neighborhood V of $p_2(K)$ in \mathbb{R}_+^n such that $\Phi(\Omega \times V) \subset \Omega \times V$. By continuity of Φ , this inclusion still holds for the closure \bar{V} of V , i.e.

$$\Phi(\Omega \times \bar{V}) \subset \Omega \times \bar{V}.$$

This inclusion implies that, for every point $(\omega, x) \in \Omega \times \bar{V}$, there exists a sequence $\{x_t\}_{t \in \mathbb{N}} \subset \bar{V}^{\mathbb{N}}$ such that $x_0 = x$, and $(\theta^{t+1}(\omega), x_{t+1}) = \Phi(\theta^t(\omega), x_t)$ for all $t \geq 0$. The sequence $\{(\theta^t(\omega), x_t)\}_{t \geq 0}$ is called a Φ -positive trajectory. Note that the first coordinate of a Φ -positive trajectory is characterized by ω and θ . Therefore a Φ -positive trajectory can be seen as a couple $(\omega, \{x_t\}_{t \geq 0})$. In order to create a past for all those Φ -positive trajectories, let us pick a point $x^* \in S \setminus (\bar{V} \cup S_0)$, and consider the product space $\mathcal{T} := \Omega \times (\bar{V} \cup \{x^*\})^{\mathbb{Z}}$ endowed with the product topology, and the homeomorphism $\Theta : \mathcal{T} \rightarrow \mathcal{T}$ defined by $\Theta(\omega, \{x_t\}_{t \in \mathbb{Z}}) = (\theta(\omega), \{x_{t+1}\}_{t \in \mathbb{Z}})$ and called the *shift operator*. Since both Ω and $\bar{V} \cup \{x^*\}$ are compact, the space \mathcal{T} is compact as well.

Every Φ -positive trajectory can be realized as an element of \mathcal{T} by creating a fixed past (i.e. $x_t = x^*$ for all $t < 0$). Then, define

$$\Gamma = \overline{\bigcup_{t \in \mathbb{Z}} \Theta^t \{ \gamma \in \mathcal{T} : \gamma \text{ is a } \Phi\text{-positive trajectory} \}}.$$

In words, Γ is the adherence in \mathcal{T} of the set of all shifted (by Θ^t for some $t \in \mathbb{Z}$) Φ -positive trajectories. Since Γ is a closed subset of the compact \mathcal{T} , it is compact as well. Moreover Γ is invariant under Θ , which implies that the restriction $\Theta|_\Gamma$ of Θ on Γ is well-defined. To simplify the presentation we still denote this restriction by Θ . The projection map $\pi_0 : \Gamma \rightarrow \Omega \times \bar{V} \cup \{x^*\}$ is defined by $\pi_0(\gamma) = (\omega, x_0)$ for all $\gamma = (\omega, \{x_t\}_t) \in \Gamma$. By definition, the map π_0 is continuous and surjective. For now on, when we write $\gamma \in \Gamma$, we mean $\gamma = (\omega, \{x_t\}_{t \in \mathbb{Z}})$.

Define the compact set of all Φ -total trajectories as

$$\Gamma_+ := \pi_0^{-1}(\Omega \times \bar{V}),$$

and the compact set of Φ -total-solution trajectory on the extinction set S_0 as

$$\Gamma_0 := \pi_0^{-1}(\Omega \times S_0).$$

The dynamic induced by Φ on $\Omega \times \bar{V}$ is linked to the dynamic induced by Θ on Γ_+ by the following semi conjugacy

$$\pi_0 \circ \Theta = \Phi \circ \pi_0. \tag{26}$$

Thus, the map Θ on Γ_+ can be seen as the extension of the map Φ on $\Omega \times \bar{V}$.

In order to write an equivalent statement of Theorem 8.5 with respect to the dynamics of Θ , we consider a subset of the invariant measures of Θ consistent with the set $\text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times S)$ in the sense of Corollary 8.8 below. For $B \subset \Gamma$ positively Θ -invariant and compact, define

$$\text{Inv}_{\mathbb{Q}}(\Theta)(B) := \{\tilde{\mu} \in \text{Inv}(\Theta)(B) : p_1^* \circ \pi_0^*(\tilde{\mu}) = \mathbb{Q}\}.$$

Proposition 8.6 *$\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)$ and $\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0)$ are compact and convex subsets of $\mathcal{P}(\Gamma)$.*

Proof Since Γ_+ and Γ_0 are positively invariant compacts, $\text{Inv}(\Theta)(\Gamma_+)$ and $\text{Inv}(\Theta)(\Gamma_0)$ are non empty, compact and convex subsets of $\mathcal{P}(\Gamma)$. Then, since $p_1^* \circ \pi_0^*$ is continuous, $\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)$ (resp. $\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0)$) is compact as closed subset of $\text{Inv}(\Theta)(\Gamma_+)$ (resp. $\text{Inv}(\Theta)(\Gamma_0)$). The convexity of $\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)$ and $\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0)$ is a consequence of the convexity of $\text{Inv}(\Theta)(\Gamma_+)$ and $\text{Inv}(\Theta)(\Gamma_0)$, and the linearity of $p_1^* \circ \pi_0^*$. \square

As a consequence of equation (26), we have

Proposition 8.7 *For every Θ -invariant measure $\tilde{\mu}$ supported on Γ_+ , $\pi_0^*(\tilde{\mu})$ is Φ -invariant.*

Proof Let $\tilde{\mu}$ be a Θ -invariant measure supported on Γ_+ . Then the measure $\pi_0^*(\tilde{\mu})$ is supported by $\Omega \times \bar{V}$. Let $B \subset \Omega \times \bar{V}$ be a Borel set. We have

$$\begin{aligned} \pi_0^*(\tilde{\mu})(\Phi^{-1}(B)) &= \tilde{\mu}(\pi_0^{-1}(\Phi^{-1}(B))) \\ &= \tilde{\mu}(\pi_0^{-1}(\Phi^{-1}(B)) \cap \Omega \times \bar{V}) \\ &= \tilde{\mu}((\Phi|_{\Omega \times \bar{V}} \circ \pi_0)^{-1}(B)) \\ &= \tilde{\mu}((\pi_0 \circ \Theta|_{\Gamma_+})^{-1}(B)) \\ &= \tilde{\mu}(\pi_0^{-1}(B)) \\ &= \pi_0^*(\tilde{\mu})(B). \end{aligned}$$

The second equality follows from the fact that the support of $\tilde{\mu}$ is included in Γ_+ , and the fourth is a consequence of the conjugacy (26). \square

Corollary 8.8 $\pi_0^*(\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+))$ is a compact and convex subset of $\text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times \overline{V})$.

Proof Since π_0^* is continuous and linear, Proposition 8.6 implies that $\pi_0^*(\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+))$ is compact and convex. Proposition 8.7 implies that $\pi_0^*(\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)) \subset \text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times \overline{V})$. \square

Remark 8.9 The definition of Θ and assumption **H3** imply that the sets Γ_0 and $\Gamma_+ \setminus \Gamma_0$ are both positively Θ -invariant. Therefore every Θ -invariant measure $\tilde{\mu}$ on Γ_+ can be written as a convex combination of two Θ -invariant measures $\tilde{\nu}_0$ and $\tilde{\nu}_1$ such that $\tilde{\nu}_0(\Gamma_0) = 1$ and $\tilde{\nu}_1(\Gamma_+ \setminus \Gamma_0) = 1$.

In order to restate Theorem 8.5 in the space of trajectories, the random matrix products (25) over Φ have to be rewritten as products over Θ . For each $i \in \{1, \dots, m\}$, define the maps $A_i : \Gamma \rightarrow \mathbf{M}_{n_i}(\mathbb{R})$ by

$$A_i(\gamma) = \begin{cases} A_i(\omega, x^*) & \text{if } x_0 = x^* \\ A_i(\omega, x_0) & \text{either.} \end{cases}$$

As (25), we write

$$A_i^t(\gamma) := A_i(\gamma) \cdots A_i(\Theta^{t-1}(\gamma)). \tag{27}$$

The conjugacy (26) implies that for all $(\omega, x) \in \Omega \times \overline{V}$ and all $\gamma \in \pi_0^{-1}(\omega, x)$, we have

$$A_i^t(\gamma) = A_i^t(\omega, x), \tag{28}$$

for all $t \geq 0$.

Then the long-term growth rates for the product (28) is

$$r_i(\gamma) := \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|A_i^t(\gamma)\|,$$

and, for a Θ -invariant measure $\tilde{\mu}$, the long-term growth rates is

$$r_i(\tilde{\mu}) = \int_{\Gamma} r_i(\gamma) d\tilde{\mu}.$$

The following proposition shows that the long-term growth rates for the product (28) defined on the trajectory space are consistent with those for the product (25) defined on $\Omega \times \overline{V}$.

Proposition 8.10 For all species i , we have

- (i) $r_i(\omega, x) = r_i(\gamma)$, for all $(\omega, x) \in \Omega \times \overline{V}$ and for all $\gamma \in \pi_0^{-1}(\omega, x)$,

(ii) for all $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)$, $\pi_0^*(\tilde{\mu}) \in \text{Inv}_{\mathbb{Q}}(\Phi)(\Omega \times \overline{V})$, and

$$r_i(\tilde{\mu}) = r_i(\pi_0^*(\tilde{\mu})).$$

Proof Assertion (i) is a consequence of equality (28), and assertion (ii) is a consequence of Corollary 8.8. □

We can now state an equivalent version of Theorem 8.5 on the space of trajectories Γ .

Theorem 8.11 *If one of the following equivalent conditions hold*

- (a) $r_*(\tilde{\mu}) := \max_{0 \leq i \leq m} r_i(\tilde{\mu}) > 0$ for every probability measure $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0)$,
or
- (b) there exist positive constants p_1, \dots, p_m such that

$$\sum_i p_i r_i(\tilde{\mu}) > 0$$

- for every ergodic probability measure $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0)$, or
- (c) there exist positive constants p_1, \dots, p_m such that

$$\sum_i p_i r_i(\omega, x) > 0$$

for every $x \in S_0$ and \mathbb{Q} -almost all $\omega \in \Omega$,

then for all $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\limsup_{t \rightarrow \infty} \Pi_t(\omega, x)(S_\eta) \leq \varepsilon \text{ for } \mathbb{Q}\text{-almost all } \omega,$$

whenever $x \in \mathbb{R}_+^n \setminus S_0$.

Remark 8.12 Condition (c) of Theorem 8.11 and (iii) Theorem 8.5 are equivalent, and the implications from conditions (iii) to (ii) and (ii) to (i) of Theorem 8.5 are direct. The proof of Theorem 8.11 (see Sect. 8.5) shows that (a), (b) and (c) of Theorem 8.11 are equivalent. Finally, condition (i) of Theorem 8.5 implies condition (a) of Theorem 8.11 as a direct consequence of assertion (ii) of Proposition 8.10. Hence, Theorems 8.5 and 8.11 are equivalent.

8.3 Random Perron-Frobenius Theorem and long-term growth rates

In this section, we first state Proposition 3.2 of Ruelle (1979a) (which we call the Random Perron-Frobenius Theorem) in its original framework, and extend it to ours. We use this extension to deduce some properties on the long-term growth rates which are crucial for the proof of Theorem 8.11. Let $\text{int } \mathbb{R}_+^d = \{x \in \mathbb{R}_+^d : \prod_i x_i > 0\}$ be the interior of \mathbb{R}_+^d .

Proposition 8.13 (Ruelle 1979a) *Let Ξ be a compact space, $\Psi : \Xi \rightarrow \Xi$ be an homeomorphism. Consider a continuous map $T : \Xi \rightarrow \mathbf{M}_d(\mathbb{R})$ and its transpose T^* defined by $T^*(\xi) = T(\xi)^*$. Write*

$$T^t(\xi) = T(\xi) \cdots T(\Psi^{t-1}\xi),$$

and assume that

A: for all $\xi \in \Xi$, $T(\xi)(\mathbb{R}_+^d) \subset \{0\} \cup \text{int } \mathbb{R}_+^d$.

Then there exist continuous maps $u, v : \Xi \rightarrow \mathbb{R}_+^d$ with $\|u(\xi)\| = \|v(\xi)\| = 1$ such that

- (i) the line bundles E (resp. F) spanned by $u(\cdot)$ (resp. $v(\cdot)$) are such that $\mathbb{R}^d = E \oplus F^\perp$ where $b \in F(\xi)^\perp$ if and only if $\langle b(\xi), v(\xi) \rangle = 0$.
- (ii) E (resp. F) is T, Ψ -invariant (resp. T^*, Ψ^{-1} -invariant), i.e. $E(\Psi(\xi)) = E(\xi)T(\xi)$ and $F(\Psi\xi)T^*(\Psi\xi) = F(\xi)$, for all $\xi \in \Xi$;
- (iii) there exist constants $\alpha < 1$ and $C > 0$ such that for all $t \geq 0$, and $\xi \in \Xi$,

$$\|bT(\xi) \cdots T(\Psi^{t-1}\xi)\| \leq C\alpha^t \|aT(\xi) \cdots T(\Psi^{t-1}\xi)\|,$$

for all unit vectors $a \in E(\xi), b \in F(\xi)^\perp$.

Our choice to call Proposition 8.13 the *Random Perron-Frobenius Theorem* is motivated by the following remark.

Remark 8.14 Assume that the map $T : \Xi \rightarrow \mathbf{M}_d(\mathbb{R})$ is constant, i.e. there exists $B \in \mathbf{M}_d(\mathbb{R})$ a positive matrix such that $T(\xi) = B$ for all $\xi \in \Xi$. Then Proposition 8.13 can be restated as follows: there exist $u, v \in \mathbb{R}_+^d$ such that $u(\xi) = u$ and $v(\xi) = v$ for all $\xi \in \Xi$; the positive vectors u and v^* are respectively the right and left eigenvector of B associated to its dominant eigenvalue (also called *Perron eigenvalue*) $r > 0$; assertion (iii) can be restated as the *strong ergodic theorem of demography*. That is

$$\lim_{t \rightarrow \infty} B^t x / r^t = v^* x u,$$

for all $x \in \text{int } \mathbb{R}_+^d$. Since $B^t x$ is the population at time t with an initial population x , the interpretation of this theorem is that the eigenvector u represents the *stable population structure*, and the coefficients of v are the *reproductive values* of the population.

In Proposition 8.13, the stable population structure and the reproductive values can not be fixed vectors whereas long-term dynamics of the population depends on the sequence of the environment incapsulated in ξ . Therefore, they have to be functions of the environment, i.e. $u, v : \Xi \rightarrow \mathbb{R}_+^d$. To interpret those functions, we look at the following consequence of assertion (iii)

$$\lim_{t \rightarrow \infty} \frac{x T^t(\Psi^{-t}\xi)}{\|x T^t(\Psi^{-t}\xi)\|} = u(\xi), \tag{29}$$

and its dual version

$$\lim_{t \rightarrow \infty} \frac{T^t(\xi)y^*}{\|T^t(\xi)y^*\|} = v(\xi)^* \tag{30}$$

The former equation appears in the proof of Proposition 8.15 as equation (31). For the sake of interpretation, assume that the environment along time has been fixed (here $\dots, \Psi^{-1}\xi, \xi, \Psi^1\xi, \dots$). Then (29) is interpreted as follows: whatever was the population a long time ago (here x), its structure today is given by $u(\xi)$. For equation (30), the interpretation is: whatever we assume to be the reproductive values in a long time (here y), the reproductive values at time $t = 0$ is given by $v(\xi)$.

In applications, the environment is represented by a stationary and ergodic process (E_t) . Here ξ represents itself a realization of this process, i.e. a possible trajectory of the environment. Therefore, there exist two stationary and ergodic processes (U_t) and (V_t) such that respectively $u(\xi)$ and $v(\xi)$ are realizations of them. Then equations (29) and (30) can be interpreted as for any initial population, in a long-term, the stage structure are given by a version of the process (U_t) and the reproductive values are given by a version of (V_t) .

Since assumption **H2** does not directly imply assumption **A** for the map $A_i(\cdot, \cdot)$, we need to extend Ruelle’s proposition to the case where

A1’: for all $\xi \in \Xi$, $T(\xi) \text{ int } \mathbb{R}_+^d \subset \text{int } \mathbb{R}_+^d$, and

A2’: there exists $s \geq 1$ such that, for all $\xi \in \Xi$, $T(\xi) \cdots T(\Psi^{s-1}\xi)(\mathbb{R}_+^d) \subset \{0\} \cup \text{int } \mathbb{R}_+^d$.

Proposition 8.15 *The conclusions of Proposition 8.13 still hold under assumptions A1’–A2’.*

Proof Define the continuous map $T' : \Xi \rightarrow \Xi \times \mathbf{M}_d(\mathbb{R})$ by

$$T'(\xi) = T(\xi) \cdots T(\Psi^{s-1}(\xi)).$$

By assumption **A2’**, $T'(\xi)\mathbb{R}_+^d \subset \{0\} \cup \text{int } \mathbb{R}_+^d$. Therefore, Proposition 8.13 applies to the map T' and to the homeomorphism Ψ^s which give us maps $u, v : \Xi \rightarrow \mathbb{R}_+^d$ with $\|u(\xi)\| = \|v(\xi)\| = 1$, their respective vector bundles $E(\cdot), F(\cdot)$, and some constants C, α verifying properties (i), (ii), and (iii).

The vector bundles $E(\cdot), F(\cdot)$ are our candidate bundles for T . We need only to check properties (ii) and (iii) for the map T as property (i) is immediate.

We claim that

$$\lim_{t \rightarrow \infty} \frac{xT^t(\Psi^{-t}\xi)}{\|xT^t(\Psi^{-t}\xi)\|} = u(\xi), \tag{31}$$

uniformly on all compact subsets of $\mathbb{R}_+^d \setminus \{0\}$. The motivation of equation (31) follows from assumption **A2’** which implies that the positive cone is contracted after every interval of time of length s . For an interpretation of (31), see Remark 8.14. Before we

prove (31), we show property (ii), i.e. $E(\cdot)$ is T, Ψ -invariant, is a consequence (31). Let $y \in \text{int } \mathbb{R}_+^d \setminus \{0\}$, and $\xi \in \Xi$. Continuity of T and equality (31) applied to y imply

$$\begin{aligned} u(\xi)T(\xi) &= \lim_{t \rightarrow \infty} \frac{yT^t(\Psi^{-t}\xi)}{\|yT^t(\Psi^{-t}\xi)\|} T(\xi) \\ &= \lim_{t \rightarrow \infty} \frac{yT(\Psi^{-t}\xi)T^t(\Psi^{-t}(\Psi\xi))}{\|yT^t(\Psi^{-t}\xi)\|} \\ &= u(\Psi\xi) \lim_{t \rightarrow \infty} \frac{\|yT(\Psi^{-t}\xi)T^t(\Psi^{-t}(\Psi\xi))\|}{\|yT^t(\Psi^{-t}\xi)\|}, \end{aligned}$$

where the final line follows from (31) with $\xi = \Psi\xi$ and $x = yT(\Psi^{-t}\xi)/\|yT(\Psi^{-t}\xi)\|$ which belongs to the compact $\{z \in \mathbb{R}_+^d : \|z\| = 1\}$ for all $t \geq 0$. This proves property (ii) for E . The same argument for the transpose T^{*} implies property (ii) for F .

Now we prove (31). Let $x \in \mathbb{R}_+^d \setminus \{0\}$ with $\|x\| = 1$. For every $t \geq 0$, define $s_t := t - [\frac{t}{s}]s$ where $[q]$ is the integer part of q . We have

$$xT^t(\Psi^{-t}\xi) = xT^{s_t}(\Psi^{-t}\xi)T^{[\frac{t}{s}]}(\Psi^{-t+s_t}\xi).$$

Since $s_t \leq s$ for all $t \geq 0$, continuity of T , and assumption **A1'** imply that there is a compact $H \subset \mathbb{R}_+^d \setminus \{0\}$ independent of x such that $xT^{s_t}(\Psi^{-t}\xi) \in H$ for all $t > 0$. Then, (31) is a consequence of inclusion (3.2) in the proof of Proposition 3.2 in Ruelle (1979a) applied to the map T' .

It remains to check property (iii): show that there exist $\alpha', C' > 0$ such that

$$\|bT^t(\xi)\| \leq C'\alpha'^t \|u(\xi)T^t(\xi)\| \quad \text{for all } t \geq s, \xi \in \Xi, b \in F(\xi)^\perp.$$

We have

$$bT^t(\xi) = bT^{s_t}(\xi)T^{[\frac{t}{s}]}(\Psi^{s_t}\xi).$$

Since $F(\cdot)$ is T^* -invariant, $bT^{s_t}(\xi) \in F(\Psi^{s_t}\xi)^\perp$ and property (iii) for T' implies

$$\frac{1}{\|bT^{s_t}(\xi)\|} \|bT^{s_t}(\xi)T^{[\frac{t}{s}]}(\Psi^{s_t}\xi)\| \leq \frac{C(\alpha^{\frac{1}{s}})^t}{\|u(\xi)T^{s_t}(\xi)\|} \|u(\xi)T^{s_t}(\xi)T^{[\frac{t}{s}]}(\Psi^{s_t}\xi)\|.$$

The continuity of T and $u(\cdot)$, and assumption **A1'** imply that there exist a constant $R \geq 0$ such that

$$\frac{\max\{\|wT^k(\xi)\| : \|w\| = 1\}}{\min\{\|u(\xi)T^k(\xi)\| : \xi \in \Xi\}} \leq R,$$

for all $k \leq s$ and all $\xi \in \Xi$. Then property (iii) is verified with $C' = CR$ and $\alpha' = \alpha^{\frac{1}{s}}$. □

Assumptions **H2-H3** imply that each continuous map $A_i : \Gamma \rightarrow \mathbf{M}_{n_i}(\mathbb{R})$ satisfies assumptions **A1'-A2'**. Hence Proposition 8.15 applies to each continuous map A_i , and to the homeomorphism Θ on the compact space Γ . Then, for each of those maps, there exist row vector maps $u_i(\cdot), v_i(\cdot)$, their respective vector bundles $E_i(\cdot), F_i(\cdot)$, and the constant $C_i, \alpha_i > 0$ satisfying properties (i), (ii), and (iii) of Proposition 8.15.

For each $i \in \{1, \dots, m\}$, define the continuous map $\xi_i : \Gamma \rightarrow \mathbb{R}$ by

$$\xi_i(\gamma) := \ln \|u_i(\gamma)A_i(\gamma)\|.$$

In the rest of this subsection, we deduce from Proposition 8.15 some crucial properties of the invasions rates.

Proposition 8.16 *For all $\gamma \in \Gamma$ and every population i , $r_i(\gamma)$ satisfies the following properties:*

(i)

$$r_i(\gamma) = \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|vA_i^t(\gamma)\|,$$

for all $v \in \mathbb{R}_+^{n_i} \setminus \{0\}$ and

(ii)

$$r_i(\gamma) = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{s=0}^{t-1} \xi_i(\Theta^s(\gamma)).$$

The proof of this proposition follows the ideas of the proof of Proposition 1 in Hofbauer and Schreiber (2010).

Proof Let $\gamma \in \Gamma$ be fixed. To prove the first part, we start by showing that

$$r_i(\gamma) = \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|u_i(\gamma)A_i^t(\gamma)\|. \tag{32}$$

Let $v \in \mathbb{R}^{n_i}, v \neq 0$. Since $\mathbb{R}^{n_i} = E_i(\gamma) \oplus F_i^\perp(\gamma)$, there exist a constant $a \in \mathbb{R}$ and a vector $w \in F_i^\perp(\gamma)$ such that $v = au_i(\gamma) + w$. Then, by Proposition 8.15, we have

$$\begin{aligned} \|vA_i^t(\gamma)\| &\leq a\|u_i(\gamma)A_i^t(\gamma)\| + \|wA_i^t(\gamma)\| \\ &\leq \|u_i(\gamma)A_i^t(\gamma)\| (a + C_i\alpha_i\|w\|). \end{aligned}$$

Hence,

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|vA_i^t(\gamma)\| \leq \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|u_i(\gamma)A_i^t(\gamma)\|$$

for all $v \in \mathbb{R}^{n_i} \setminus \{0\}$. Since $\|A_i^t(\gamma)\| = \sup_{\|v\|=1} \|vA_i^t(\gamma)\|$, the last inequality implies that

$$r_i(\gamma) \leq \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|u_i(\gamma)A_i^t(\gamma)\| \leq r_i(\gamma),$$

which proves the equality (32).

Now, we consider positive vector $v \in \mathbb{R}_+^{n_i} \setminus \{0\}$. We show that the equality (32) is also satisfied for v . We write $v = au_i(\gamma) + w$ with $a > 0$ and $w \in F_i^\perp(\gamma)$. Proposition 8.15 implies

$$\begin{aligned} \|vA_i^t(\gamma)\| &\geq a\|u_i(\gamma)A_i^t(\gamma)\| - \|wA_i^t(\gamma)\| \\ &\geq \|u_i(\gamma)A_i^t(\gamma)\| (a - C_i\alpha_i^t\|w\|). \end{aligned}$$

Since $a > 0$,

$$r_i(\gamma) \geq \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|vA_i^t(\gamma)\| \geq \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|u_i(\gamma)A_i^t(\gamma)\| = r_i(\gamma),$$

which completes the proof of assertion (i).

The second assertion results directly from the first assertion and the following equalities:

$$\begin{aligned} \ln \|u_i(\gamma)A_i^{t+1}(\gamma)\| &= \ln \|u_i(\gamma)A_i^t(\gamma)A_i(\Theta^t(\gamma))\| \\ &= \ln \|u_i(\Theta(\gamma)^t)A_i(\Theta^t(\gamma))\| + \ln \|u_i(\gamma)A_i^t(\gamma)\| \\ &= \xi_i(\Theta^t(\gamma)) + \ln \|u_i(\gamma)A_i^t(\gamma)\|. \end{aligned}$$

The second step is a consequence of the invariance of the line bundle E_i . □

Recall that $\Gamma_+ = \pi_0^{-1}(\Omega \times \bar{V})$ and $\Gamma_0 = \pi_0^{-1}(\Omega \times S_0)$.

Corollary 8.17 *For all $\gamma \in \Gamma_+ \setminus \Gamma_0$, and every $i \in \{1, \dots, m\}$,*

$$r_i(\gamma) \leq 0.$$

Proof Fix $i \in \{1, \dots, m\}$, and $\gamma \in \Gamma_+ \setminus \Gamma_0$ with $(\omega, x) := \pi_0(\gamma)$. By definition of $\Gamma_+ \setminus \Gamma_0$, $x^i \in \mathbb{R}_+^{n_i}$ and $x^i \neq 0$. We have

$$\begin{aligned} x^i A_i^t(\gamma) &= x^i A_i(\gamma) \cdots A_i(\Theta^{t-1}(\gamma)) \\ &= x^i A_i(\omega, x) \cdots A_i(\Phi^{t-1}(\omega, x)) \\ &= p_2(\Phi^t(\omega, x)), \end{aligned}$$

where the second equality is a consequence of (28), and the third one follows from the definition of the cocycle Φ . Assumption **H4'** implies that there exists $T > 0$ such that $p_2(\Phi^t(\omega, x))$ belongs to the compact set \bar{V} for all $t \geq T$, which implies that there exists $R > 0$ such that $\|x^i A_i^t(\gamma)\| \leq R$ for all $t \geq T$. Assertion (i) of Proposition 8.16 applied to $v = x^i$ concludes the proof. □

Now we give some properties of the invasion rate with respect to a Θ -invariant probability measure.

Proposition 8.18 *The invasion rate of each population i with respect to an Θ -invariant measure $\tilde{\mu}$ satisfies the following property:*

$$r_i(\tilde{\mu}) = \int_{\Gamma} \zeta_i(\gamma) d\tilde{\mu}.$$

Proof This result is a direct consequence of property (ii) of Proposition 8.16 and the Birkhoff’s Ergodic Theorem applied to the continuous maps Θ and ζ . \square

Proposition 8.19 *Let $\tilde{\mu}$ be a Θ -invariant measure. If $\tilde{\mu}$ is supported by $\Gamma_+ \setminus \Gamma_0$, then $r_i(\tilde{\mu}) = 0$ for all $i \in \{1, \dots, m\}$.*

Proof Let $\tilde{\mu}$ be such a probability measure. Fix $i \in \{1, \dots, m\}$, and define the set $\Gamma^{i,\eta} := \{\gamma \in \Gamma_+ : \|p_2(\pi_0(\Theta^t(\gamma)))\|^i > \eta\}$. By assumption on the measure $\tilde{\mu}$, there exists a real number $\eta^* > 0$ such that $\tilde{\mu}(\Gamma^{i,\eta}) > 0$ for all $\eta < \eta^*$.

The Poincaré recurrence theorem applies to the map Θ , and implies that for each $\eta < \eta^*$,

$$\tilde{\mu}(\{\gamma \in \Gamma^{i,\eta} \mid \Theta^t(\gamma) \in \Gamma^{i,\eta} \text{ infinitely often}\}) = 1. \tag{33}$$

Recall that the conjugacy (26) implies that for every $\gamma \in \Gamma_+$ with $\pi_0(\gamma) = (\omega, x) \in \Omega \times \bar{V} \setminus S_0$, we have

$$\begin{aligned} p_2(\pi_0(\Theta^t(\gamma)))^i &= p_2(\Phi^t(\pi_0(\gamma)))^i \\ &= x^i A_i^t(\gamma). \end{aligned}$$

Then, equality (33) means that for $\tilde{\mu}$ -almost all $\gamma \in \Gamma^{i,\eta}$ with $0 < \eta < \eta^*$, $\|x^i A_i^t(\gamma)\| > \eta$ infinitely often. Therefore, Proposition 8.16 (i), applied to $v = x^i$, implies that $r_i(\gamma) = \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|x^i A_i^t(\gamma)\| \geq 0$ for $\tilde{\mu}$ -almost all $\gamma \in \Gamma^{i,\eta}$, with $\eta < \eta^*$. Hence $r_i(\gamma) \geq 0$ for $\tilde{\mu}$ -almost all $\gamma \in \bigcup_{\eta \geq \frac{1}{\eta^*}} \Gamma^{i,1/\eta} = \Gamma_+ \setminus \Gamma_0$. Corollary 8.17 completes the proof. \square

8.4 Properties of the empirical occupation measures

Given a trajectory $\gamma \in \Gamma_+$, the *empirical occupation measure* at time $t \in \mathbb{N}$ of $\{\Theta^s(\gamma)\}_{s \geq 0}$ is

$$\tilde{\Lambda}_t(\gamma) := \frac{1}{t} \sum_{s=0}^{t-1} \delta_{\Theta^s(\gamma)},$$

and given a point $(\omega, x) \in \Omega \times \bar{V}$, the empirical occupation measure at time $t \in \mathbb{N}$ of $\{\Phi^s(\omega, x)\}_{s \geq 0}$ is

$$\Lambda_t(\omega, x) := \frac{1}{t} \sum_{s=0}^{t-1} \delta_{\Phi^s(\omega, x)}.$$

In this way, $\Lambda_t(\omega, x)(\Omega \times B) = \Pi_t(\omega, x)(B)$ for every Borel subset $B \subset \bar{V}$, and $x \in \bar{V}$.

The dynamics Θ and Φ being semi-conjugated by π_0 , their respective empirical occupation measures are semi-conjugated by π_0^* as follows.

Lemma 8.20 *Let $\gamma \in \Gamma_+$. Then for all $t \geq 0$ we have*

$$\pi_0^*(\tilde{\Lambda}_t(\gamma)) = \Lambda_t(\pi_0(\gamma)).$$

Proof Let $B \subset \Omega \times \bar{V}$ be a Borel set, and $\gamma \in \Gamma_+$. Then we have

$$\begin{aligned} \pi_0^*(\tilde{\Lambda}_t(\gamma))(B) &= \tilde{\Lambda}_t(\gamma)(\pi_0^{-1}(B)) \\ &= \frac{1}{t} \sum_{s=0}^{t-1} \delta_{\Theta^s(\gamma)}(\pi_0^{-1}(B)) \\ &= \frac{1}{t} \sum_{s=0}^{t-1} \delta_{\Phi^s(\pi_0(\gamma))}(B) \\ &= \Lambda_t(\pi_0(\gamma))(B). \end{aligned}$$

The third equality is a consequence of the semi conjugacy (26). □

Proposition 8.21 *There exists $\tilde{\Omega}$ with $\mathbb{Q}(\tilde{\Omega}) = 1$ such that for all $\gamma \in \pi_0^{-1}(\tilde{\Omega} \times \bar{V})$, the set of all weak* limit point of the family of probability measures $\{\tilde{\Lambda}_t(\gamma)\}_{t \in \mathbb{N}}$ is a non-empty subset of $\text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)$.*

Proof Since \mathbb{Q} is ergodic (assumption **H4'**), Birkhoff's Ergodic Theorem implies that there exists a subset $\tilde{\Omega} \subset \Omega$ such that $\mathbb{Q}(\tilde{\Omega}) = 1$, and for all $\omega \in \tilde{\Omega}$,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=0}^{t-1} \delta_{\theta^s(\omega)} = \mathbb{Q} \tag{34}$$

(in the weak* topology). Let $(\omega, x) \in \tilde{\Omega} \times \bar{V}$ and $\gamma \in \pi_0^{-1}(\omega, x) \subset \Gamma_+$. For all $t \in \mathbb{N}$, we have

$$p_1^* \circ \pi_0^*(\tilde{\Lambda}_t(\gamma)) = \frac{1}{t} \sum_{s=0}^{t-1} \delta_{\theta^s(\omega)}. \tag{35}$$

Since Γ_+ is positively Θ -invariant and compact, the set of all weak* limit point of the family of probability measures $\{\tilde{\Lambda}_t(\gamma)\}_{t \in \mathbb{N}}$ is a non-empty subset of $\mathcal{P}(\Gamma_+)$. Since the maps p_1 and π_0 are continuous, equalities (34) and (35) imply that $p_1^* \circ \pi_0^*(\tilde{\mu}) = \mathbb{Q}$. Moreover, Theorem 6.9 in Walters (1982) implies that $\tilde{\mu}$ is Θ -invariant. Therefore, $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)$, which concludes the proof. \square

Recall that $S_\eta = \{x \in S : \|x^i\| \leq \eta \text{ for some } i\}$, and define the subset $\Gamma_\eta := \pi_0^{-1}(\Omega \times S_\eta)$.

Proposition 8.22 *If conation (a) of Theorem 8.11 is satisfied, then for all $\varepsilon > 0$ there exists $\eta^* > 0$ such that*

$$\tilde{\mu}(\Gamma_\eta) < \varepsilon,$$

for all $\eta < \eta^*$ and all $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+ \setminus \Gamma_0)$.

Proof If false, there exist $\varepsilon > 0$ and a sequence of measures $\{\tilde{\mu}_n\}_{n \in \mathbb{N}} \subset \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+ \setminus \Gamma_0)$ such that $\tilde{\mu}_n(\Gamma_{1/n}) > \varepsilon$ for all $n \geq 1$. By Proposition 8.6, let $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)$ be a weak* limit point of the sequence $\{\tilde{\mu}_n\}_{n \in \mathbb{N}}$. Proposition 8.19 implies that $r_*(\tilde{\mu}_n) = 0$ for all $n \geq 0$. Proposition 8.18 and weak* convergence imply that $0 = \lim_{n \rightarrow \infty} r_i(\tilde{\mu}_n) = r_i(\tilde{\mu})$ for all i . Hence, $r_*(\tilde{\mu}) = 0$. The Portmanteau theorem (see e.g. Theorem 2.1. in Billingsley (1999)) applied to the closed set $\Gamma_{1/n}$ implies that for all $n \geq 1$,

$$\begin{aligned} \tilde{\mu}(\Gamma_{1/n}) &\geq \liminf_{m \rightarrow \infty} \tilde{\mu}_m(\Gamma_{1/n}) \\ &\geq \liminf_{m \rightarrow \infty} \tilde{\mu}_m(\Gamma_{1/m}) \\ &\geq \varepsilon. \end{aligned}$$

Therefore $\tilde{\mu}(\Gamma_0) = \tilde{\mu}(\cap_n \Gamma_{1/n}) \geq \varepsilon$. Remark 8.9 implies there exist $\alpha > 0$ such that $\tilde{\mu} = \alpha \tilde{\nu}_0 + (1 - \alpha) \tilde{\nu}_1$ where $\tilde{\nu}_j$ are Θ -invariant probability measures satisfying $\tilde{\nu}_0(\Gamma_0) = 1$ and $\tilde{\nu}_1(\Gamma_+ \setminus \Gamma_0) = 1$. By Proposition 8.19, $r_i(\tilde{\nu}_1) = 0$ for all $i \in \{1, \dots, k\}$. Condition (a) implies that $r_*(\tilde{\nu}_0) > 0$, in which case $0 = r_*(\tilde{\mu}) = \alpha r_*(\tilde{\nu}_0) > 0$ which is a contradiction. \square

8.5 Proof of Theorem 8.11

First, we show that condition (a) of Theorem 8.11 implies that for all $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\limsup_{t \rightarrow \infty} \Pi_t(\omega, x)(S_\eta) \leq \varepsilon \quad \text{for } \mathbb{Q}\text{-almost all } \omega,$$

whenever $x \in \mathbb{R}_+^n \setminus S_0$. Second, we prove the equivalence of conditions (a), (b) and (c).

Let $\tilde{\Omega} \subset \Omega$ be defined as in Proposition 8.21. Choose $(\omega', x') \in \tilde{\Omega} \times \mathbb{R}_+^n \setminus S_0$. By definition of the set \bar{V} , there exists a time $T \geq 0$ such that $\Phi^t(\omega', x') \in \Omega \times \bar{V}$, for

all $t \geq T$. Choose $\gamma \in \pi_0^{-1}(\Phi^T(\omega', x')) \subset \Gamma_+ \setminus \Gamma_0$. Since μ is a weak* limit point of the family $\{\Lambda_t(\Phi^T(\omega', x'))\}_{t \geq 0}$ if and only if it is a weak* limit point of the family $\{\Lambda_t(\omega', x')\}_{t \geq 0}$, we do not loss generality by considering $\{\Lambda_t(\Phi^T(\omega', x'))\}_{t \geq 0}$. Since $\Omega \times \bar{V}$ is compact, the set of all weak* limit points of the family of probability measures $\{\Lambda_t(\Phi^T(\omega', x'))\}_{t \in \mathbb{N}}$ is a non-empty subset of $\mathcal{P}(\Omega \times \bar{V})$. Let $\mu = \lim_{k \rightarrow \infty} \Lambda_{t_k}(\omega, x)$ be such a weak* limit point. Since Γ_+ is positively Θ -invariant and compact, passing to a subsequence if necessary, there exists $\tilde{\mu} = \lim_{k \rightarrow \infty} \tilde{\Lambda}_{t_k}(\gamma) \in \mathcal{P}(\Gamma_+)$. By Proposition 8.21, $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_+)$. Furthermore by Lemma 8.20 and continuity of π_0 , $\pi_0^*(\tilde{\mu}) = \mu$. Hence, Proposition 8.18, the continuity of the map ζ , and property (ii) of Proposition 8.16, imply the following equalities for all i :

$$\begin{aligned} r_i(\tilde{\mu}) &= \int_{\Gamma} \zeta(\eta) d\tilde{\mu}(\eta) \\ &= \lim_{k \rightarrow \infty} \frac{1}{t_k} \sum_{s=0}^{t_k-1} \zeta(\Theta^s(\gamma)) \\ &\leq r_i(\gamma). \end{aligned}$$

Hence, by Corollary 8.17,

$$r_i(\tilde{\mu}) \leq 0, \text{ for all } i.$$

Remark 8.9 implies there exists $\alpha \geq 0$ such that $\tilde{\mu} = \alpha \tilde{\nu}_0 + (1 - \alpha) \tilde{\nu}_1$ where $\tilde{\nu}_j$ are invariant probability measure satisfying $\tilde{\nu}_0(\Gamma_0) = 1$ and $\tilde{\nu}_1(\Gamma_+ \setminus \Gamma_0) = 1$. By Proposition 8.19, $r_i(\tilde{\nu}_1) = 0$ for all $i \in \{1, \dots, k\}$. Condition (a) implies $r_*(\tilde{\nu}_0) > 0$. Therefore α must be zero, i.e. $\tilde{\mu}(\Gamma_+ \setminus \Gamma_0) = 1$. Fix $\varepsilon > 0$. By Proposition 8.22 there exists $\eta^* > 0$ such that

$$\tilde{\mu}(\Gamma_\eta) < \varepsilon, \quad \forall \eta < \eta^*,$$

which implies

$$\mu(\Omega \times S_\eta) < \varepsilon, \quad \forall \eta < \eta^*.$$

Since η^* does not depend on μ , we have

$$\limsup_{t \rightarrow \infty} \Lambda_t(\omega', x')(\Omega \times S_\eta) < \varepsilon, \quad \forall \eta < \eta^*,$$

for all $x' \in \mathbb{R}_+^n \setminus S_0$ and $\omega' \in \tilde{\Omega}$, which concludes the first part of the proof.

Next, we show the equivalence of conditions (a) and (b). We need the following version of the minimax theorem (see, e.g., Simmons 1998):

Theorem 8.23 (Minimax theorem) *Let A, B be Hausdorff topological vector spaces and let $\mathcal{L} : A \times B \rightarrow \mathbb{R}$ be a continuous bilinear function. Finally, let E and F be nonempty, convex, compact subsets of A and B , respectively. Then*

$$\min_{a \in E} \max_{b \in F} \mathcal{L}(a, b) = \max_{b \in F} \min_{a \in E} \mathcal{L}(a, b).$$

We have that

$$\min_{\tilde{\mu}} \max_i r_i(\tilde{\mu}) = \min_{\tilde{\mu}} \max_p \sum_i p_i r_i(\tilde{\mu})$$

where the minimum is taken over $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0)$ and the maximum over $p \in \Delta := \{p \in \mathbb{R}_+^m : \sum_i p_i = 1\}$. Define A to be the dual space to the space of bounded continuous functions from Γ_0 to \mathbb{R} and define $B = \mathbb{R}^m$. Let $E = \Delta$, and $D = \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0) \subset A$ which is nonempty, convex and compact by Proposition 8.6. Let $\mathcal{L} : A \times B \rightarrow \mathbb{R}$ the bilinear function defined by $\mathcal{L}(\tilde{\mu}, p) := \sum_i p_i r_i(\tilde{\mu})$. Proposition 8.18 implies that \mathcal{L} is continuous. With these choices, the Minimax theorem implies that

$$\min_{\tilde{\mu}} \max_i r_i(\tilde{\mu}) = \max_{p \in \Delta} \min_{\tilde{\mu}} \sum_i p_i r_i(\tilde{\mu}) \tag{36}$$

where the minimum is taken over $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0)$. By the ergodic decomposition theorem for random dynamical systems (see Lemma 6.19 in Crauel (2002)), the minimum of the right hand side of (36) is attained at an ergodic probability measure with support in Γ_0 . Thus, the equivalence of the conditions is established.

Finally, we show the equivalence of condition (b) and (c). As a direct consequence of assertion (i) of Proposition 8.10, condition (c) implies (b). To prove the other direction, let $\tilde{\Omega} \subset \Omega$ be defined as in the proof of Proposition 8.21. Choose $(\omega', x') \in \tilde{\Omega} \times S_0$. By the same arguments as above, there exist $T > 0, \gamma \in \pi_0^{-1}(\Phi^T(\omega', x')) \subset \Gamma_0$ and $\tilde{\mu} \in \text{Inv}_{\mathbb{Q}}(\Theta)(\Gamma_0)$ such that

$$\begin{aligned} r_i(\tilde{\mu}) &= \int_{\Gamma} \xi(\eta) d\tilde{\mu}(\eta) \\ &= \lim_{k \rightarrow \infty} \frac{1}{t_k} \sum_{s=0}^{t_k-1} \xi(\Theta^s(\gamma)) \\ &\leq r_i(\gamma). \end{aligned}$$

Assertion (i) of Proposition 8.10 implies that $r_i(\gamma) = r_i(\Phi^T(\omega', x'))$. Since $\Phi^T(\omega', x')$ is on the same trajectory that (ω', x') , $r_i(\tilde{\mu}) \leq r_i(\Phi^T(\omega', x')) = r_i(\omega', x')$. Writing $\tilde{\mu}$ as a convex combination of ergodic probability measures, condition (b) implies $\sum_i p_i r_i(\omega', x') > 0$. □

Acknowledgments GR was supported by the Swiss National Science Foundation Grant 137273 and a start-up grant to SJS from the College of Biological Sciences, University of California, Davis. SJS was supported in part by U.S. National Science Foundation Grants EF-0928987 and DMS-1022639. The authors thank two anonymous referees for their very useful comments on an earlier draft of this paper.

References

- Allesina S, Levine JM (2011) A competitive network theory of species diversity. *Proc Natl Acad Sci* 108:5638
- Amarasekare P (2003) Competitive coexistence in spatially structured environments: a synthesis. *Ecol Lett* 6:1109–1122
- Armstrong RA, McGehee R (1980) Competitive exclusion. *Am Nat* 116:151–170
- Arnold L (1998) Random dynamical systems. In: Springer monographs in mathematics, Springer, Berlin
- Bascompte J, Possingham H, Roughgarden J (2002) Patchy populations in stochastic environments: critical number of patches for persistence. *Am Nat* 159:128–137
- Benaïm M, Hofbauer J, Sandholm W (2008) Robust permanence and impermanence for the stochastic replicator dynamics. *J Biol Dyn* 2:180–195
- Benaïm M, Schreiber SJ (2009) Persistence of structured populations in random environments. *Theoret Popul Biol* 76:19–34
- Bhattacharya R, Majumdar M (2007) Random dynamical systems. Cambridge University Press, Cambridge
- Billingsley P (1999) Convergence of probability measures. In: Wiley series in probability and statistics: probability and statistics, 2nd edn. Wiley, New York
- Boyce MS, Haridas CV, Lee CT, the NCEAS Stochastic Demography Working Group (2006) Demography in an increasingly variable world. *Trends Ecol Evol* 21:141–148
- Buss L, Jackson J (1979) Competitive networks: nontransitive competitive relationships in cryptic coral reef environments. *Am Nat* 113:223–234
- Butler GJ, Waltman P (1986) Persistence in dynamical systems. *J Differ Equ* 63:255–263
- Cameron DD, White A, Antonovics J (2009) Parasite-grass-forb interactions and rock-paper-scissor dynamics: predicting the effects of the parasitic plant *Rhinanthus minor* on host plant communities. *J Ecol* 97:1311–1319
- Cantrell RS, Cosner C (2003) Spatial ecology via reaction-diffusion equations. In: Wiley series in mathematical and computational biology. Wiley, Chichester
- Caswell H (2001) Matrix population models. Sinauer, Sunderland
- Chesson PL (1982) The stabilizing effect of a random environment. *J Math Biol* 15:1–36
- Chesson PL (1985) Coexistence of competitors in spatially and temporally varying environments: a look at the combined effects of different sorts of variability. *Theoret Popul Biol* 28:263–287
- Chesson PL (1994) Multispecies competition in variable environments. *Theoret Popul Biol* 45:227–276
- Chesson PL (2000a) General theory of competitive coexistence in spatially-varying environments. *Theoret Popul Biol* 58:211–237
- Chesson PL (2000b) Mechanisms of maintenance of species diversity. *Annu Rev Ecol Syst* 31:343–366
- Chesson PL, Warner RR (1981) Environmental variability promotes coexistence in lottery competitive systems. *Am Nat* 117:923
- Collie JS, Spencer PD (1994) Modeling predator-prey dynamics in a fluctuating environment. *Can J Fish Aquat Sci* 51(12):2665–2672
- Costantino RF, Cushing JM, Dennis B, Desharnais RA (1995) Experimentally induced transitions in the dynamic behaviour of insect populations. *Nature* 375:227–230
- Costantino RF, Cushing JM, Dennis B, Desharnais RA (1997) Chaotic dynamics in an insect population. *Science* 275:389–391
- Crauel H (2002) Random probability measures on Polish spaces. In: Stochastics monographs, vol 11. Taylor & Francis, London
- Cushing JM (1998) An introduction to structured population dynamics. In: CBMS-NSF Regional Conference Series in Applied Mathematics, vol 71. Society for Industrial and Applied Mathematics (SIAM), Philadelphia
- Davies KF, Chesson P, Harrison S, Inouye BD, Melbourne BA, Rice KJ (2005) Spatial heterogeneity explains the scale dependence of the native-exotic diversity relationships. *Ecology* 86:1602–1610
- Dennis B, Desharnais RA, Cushing JM, Costantino RF (1995) Nonlinear demographic dynamics: mathematical models, statistical methods, and biological experiments. *Ecol Monogr* 65:261–282
- Evans SN, Ralph P, Schreiber SJ, Sen A (2013) Stochastic growth rates in spatio-temporal heterogeneous environments. *J Math Biol* 66:423–476
- Freedman HI, Waltman P (1977) Mathematical analysis of some three species food chains. *Math Biosci* 68:213–231

- Garay BM, Hofbauer J (2003) Robust permanence for ecological differential equations, minimax, and discretizations. *SIAM J Math Anal* 34:1007–1039
- Gonzalez A, Holt RD (2002) The inflationary effects of environmental fluctuations in source-sink systems. *Proc Natl Acad Sci* 99:14872–14877
- Gyllenberg M, Osipov AV, Söderbacka G (1996) Bifurcation analysis of a metapopulation model with sources and sinks. *J Nonlinear Sci* 6:329–366
- Hardin DP, Takáč P, Webb GF (1988) Asymptotic properties of a continuous-space discrete-time population model in a random environment. *J Math Biol* 26:361–374
- Harrison S, Quinn JF (1989) Correlated environments and the persistence of metapopulations. *Oikos* 56(3):293–298
- Hastings A (1988) Food web theory and stability. *Ecology* 69:1665–1668
- Hastings A, Botsford LW (2006) Persistence of spatial populations depends on returning home. *Proc Natl Acad Sci* 103:6067–6072
- Henson SM, Cushing JM (1997) The effect of periodic habitat fluctuations on a nonlinear insect population model. *J Math Biol* 36:201–226
- Hofbauer J (1981) A general cooperation theorem for hypercycles. *Monatshefte für Mathematik* 91:233–240
- Hofbauer J, Schreiber SJ (2010) Robust permanence for interacting structured populations. *J Differ Equ* 248:1955–1971
- Hofbauer J, Sigmund K (1998) *Evolutionary games and population dynamics*. Cambridge University Press, Cambridge
- Hofbauer J, So JWH (1989) Uniform persistence and repellers for maps. *Proc Am Math Soc* 107:1137–1142
- Holyoak M, Leibold MA, Holt RD (eds) (2005) *Spatial dynamics and ecological communities*. University of Chicago Press, Metacommunities
- Hutson V (1984) A theorem on average Liapunov functions. *Monatsh Math* 98:267–275
- Hutson V, Moran W (1987) Repellers in reaction-diffusion systems. *Rocky Mt J Math* 17(3):301–314
- Hutson V, Schmitt K (1992) Permanence and the dynamics of biological systems. *Math Biosci* 111:1–71
- Jansen VAA, Sigmund K (1998) Shaken not stirred: on permanence in ecological communities. *Theoret Popul Biol* 54:195–201
- Jansen VAA, Yoshimura J (1998) Populations can persist in an environment consisting of sink habitats only. *Proc Natl Acad Sci USA* 95:3696–3698
- Jin Y, Zhao XQ (2009) Spatial dynamics of a nonlocal periodic reaction-diffusion model with stage structure. *SIAM J Math Anal* 40(6):2496–2516
- Johst K, Wissel C (1997) Extinction risk in a temporally correlated fluctuating environment. *Theoret Popul Biol* 52:91–100
- Kerr B, Riley M, Feldman M, Bohannan J (2002) Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* 418:171–174
- Kirkup B, Riley M (2004) Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissors in vivo. *Nature* 428:412–414
- Kon R, Saito Y, Takeuchi Y (2004) Permanence of single-species stage-structured models. *J Math Biol* 48:515–528
- Kuang JJ, Chesson PL (2009) Coexistence of annual plants: Generalist seed predation weakens the storage effect. *Ecology* 90:170–182
- Lankau RA, Strauss SY (2007) Mutual feedbacks maintain genetic diversity and species diversity in a plant community. *Science* 317:1561–1563
- Loreau M, Mouquet N, Gonzalez A (2003) Biodiversity as spatial insurance in heterogeneous landscapes. *Proc Natl Acad Sci* 100:12765–12770
- Lundberg P, Ranta E, Ripa J, Kaitala V (2000) Population variability in space and time. *Trends Ecol Evol* 15:460–464
- Magal P, McCluskey CC, Webb GF (2010) Lyapunov functional and global asymptotic stability for an infection-age model. *Appl Anal Int J* 89:1109–1140
- Mandelbrot BB (1982) *The fractal geometry of nature*. Times Books, New York
- Matthews DP, Gonzalez A (2007) The inflationary effects of environmental fluctuations ensure the persistence of sink metapopulations. *Ecology* 88:2848–2856
- May RM (1975) *Stability and complexity in model ecosystems*, 2nd edn. Princeton University Press, Princeton
- May RM, Hassell MP (1981) The dynamics of multiparasitoid-host interactions. *Am Nat* 117:234–261

- May RM, Leonard W (1975) Nonlinear aspects of competition between three species. *SIAM J Appl Math* 29:243–252
- May RM (1973) Stability in randomly fluctuating versus deterministic environments. *Am Nat* 1973:621–650
- McGehee R, Armstrong RA (1977) Some mathematical problems concerning the ecological principle of competitive exclusion. *J Differ Equ* 23:30–52
- Metz JAJ, de Jong TJ, Klinkhamer PGL (1983) What are the advantages of dispersing; a paper by Kuno extended. *Oecologia* 57:166–169
- Metz JAJ, Gyllenberg M (2001) How should we define fitness in structured metapopulation models? including an application to the calculation of evolutionarily stable dispersal strategies. *Proc R Soc B* 268:499–508
- Metz JAJ, Diekmann O (1986) The dynamics of physiologically structured populations. In: *Lecture notes in biomathematics*, vol 68. Springer, Berlin
- Mouquet N, Loreau M (2003) Community patterns in source-sink metacommunities. *Am Nat* 162:544–557
- Nahum JR, Harding BN, Kerr B (2011) Evolution of restraint in a structured rock-paper-scissors community. *Proc Natl Acad Sci* 108:10831–10838
- Paquin C, Adams J (1983) Relative fitness can decrease in evolving asexual populations of *S. cerevisiae*. *Nature* 306:368–371
- Petchey OL, Gonzalez A, Wilson HB (1997) Effects on population persistence: the interaction between environmental noise colour, intraspecific competition and space. *Proc Biol Sci* 264(1389):1841–1847
- Petchey OL (2000) Environmental colour affects aspects of single-species population dynamics. *Proc R Soc Lond Ser B Biol Sci* 267:747–754
- Pollicott M (2010) Maximal lyapunov exponents for random matrix products. *Invent Math* 181:209–226
- Ripa J, Ives AR (2003) Food web dynamics in correlated and autocorrelated environments. *Theoret Popul Biol* 64:369–384
- Rokhlin VA (1964) Exact endomorphism of a Lebesgue space. *Izv Acad Sci USSR, Ser Mat* 25:499–530
- Roughgarden J (1975) A simple model for population dynamics in stochastic environments. *Am Nat* 109(970):713–736
- Roughgarden J (1979) *Theory of population genetics and evolutionary ecology*. Macmillan, New York
- Roy M, Holt RD, Barfield M (2005) Temporal autocorrelation can enhance the persistence and abundance of metapopulations comprised of coupled sinks. *Am Nat* 166:246–261
- Ruelle D (1979a) Analyticity properties of the characteristic exponents of random matrix products. *Adv Math* 32(1):68–80. ISSN 0001–8708
- Ruelle D (1979b) Ergodic theory of differentiable dynamical systems. *IHES Publ Math* 50:27–58
- Salceanu PL, Smith HL (2009a) Lyapunov exponents and persistence in some discrete dynamical systems. *Discrete Contin Dyn Syst* 12:187–203
- Salceanu PL, Smith HL (2009b) Lyapunov exponents and uniform weak normally repelling invariant sets. In: *Positive systems*. Springer, Berlin
- Salceanu PL, Smith HL (2010) Persistence in a discrete-time, stage-structured epidemic model. *J Differ Equ Appl* 16(1):73–103
- Schmidt KA (2004) Site fidelity in temporally correlated environments enhances population persistence. *Ecol Lett* 7:176–184
- Schreiber SJ (1997) Generalist and specialist predators that mediate permanence in ecological communities. *J Math Biol* 36:133–148
- Schreiber SJ (2000) Criteria for C^r robust permanence. *J Differ Equ* 162:400–426
- Schreiber SJ (2004) Coexistence for species sharing a predator. *J Differ Equ* 196:209–225
- Schreiber SJ (2006) Persistence despite perturbations for interacting populations. *J Theoret Biol* 242:844–852
- Schreiber SJ (2007) On persistence and extinction of randomly perturbed dynamical systems. *Discrete Contin Dyn Syst B* 7:457–463
- Schreiber SJ, Benaïm M, Atchadé KAS (2011) Persistence in fluctuating environments. *J Math Biol* 62:655–683
- Schreiber SJ, Rittenhouse S (2004) From simple rules to cycling in community assembly. *Oikos* 105:349–358
- Schreiber SJ (2010) Interactive effects of temporal correlations, spatial heterogeneity, and dispersal on population persistence. *Proc R Soc Biol Sci* 277:1907–1914
- Schreiber SJ (2012) Persistence for stochastic difference equations: a mini-review. *J Differ Equ Appl* 18:1381–1403

- Schreiber SJ, Killingback TP (2013) Cycling in space: Persistence of rock-paper-scissor metacommunities. *Theoret Popul Biol* 86:1–11
- Schuster P, Sigmund K, Wolff R (1979) Dynamical systems under constant organization 3: Cooperative and competitive behavior of hypercycles. *J Differ Equ* 32:357–368
- Simmons S (1998) *Minimax and monotonicity*. Springer, Berlin
- Sinervo B, Lively C (1996) The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* 380:240–243
- Smith HL, Thieme HR (2011) Dynamical systems and population persistence. In: *Graduate studies in mathematics*, vol 118. American Mathematical Society, Providence
- Snyder RE (2007) Spatiotemporal population distributions and their implications for species coexistence in a variable environment. *Theoret Popul Biol* 72(1):7–20
- Snyder RE (2008) When does environmental variation most influence species coexistence? *Theoret Ecol* 1(3):129–139
- Snyder RE, Chesson PL (2003) Local dispersal can facilitate coexistence in the presence of permanent spatial heterogeneity. *Ecol Lett* 6(4):301–309
- Thieme HR (2009) Spectral bound and reproduction number for infinite-dimensional population structure and time heterogeneity. *SIAM J Appl Math* 70:188–211
- Thieme HR (2011) Global stability of the endemic equilibrium in infinite dimension: Lyapunov functions and positive operators. *J Differ Equ* 250:3772–3801
- Tuljapurkar S (1990) *Population dynamics in variable environments*. Springer, New York
- Tuljapurkar S, Coulson T, Steiner UK (2012) Structured population models: introduction. *Theoret Popul Biol* 82:241–243
- Tuljapurkar S, Haridas CV (2006) Temporal autocorrelation and stochastic population growth. *Ecol Lett* 9:327–337
- Vandermeer J, Pascual M (2005) Competitive coexistence through intermediate polyphagy. *Ecol Complex* 3:37–43
- Vasseur DA, Yodzis P (2004) The color of environmental noise. *Ecology* 85:1146–1152
- Walters P (1982) An introduction to ergodic theory. In: *Graduate texts in mathematics*, vol 79. Springer, New York
- Xu D, Zhao XQ (2003) A nonlocal reaction-diffusion population model with stage structure. *Can Appl Math Q* 11:303–319
- Zhao XQ, Hutson V (1994) Permanence in kolmogorov periodic predator-prey models with diffusion. *Nonlinear Anal Theory Methods Appl* 23(5):651–668