Mathematical Biology

Persistence in fluctuating environments

**Sebastian J. Schreiber · Michel Benaïm ·
Kolawolé A. S. Atchadé**

Abstract Understanding under what conditions interacting populations, whether they be plants, animals, or viral particles, coexist is a question of theoretical and practical importance in population biology. Both biotic interactions and environmental fluctuations are key factors that can facilitate or disrupt coexistence. To better understand this interplay between these deterministic and stochastic forces, we develop a mathematical theory extending the nonlinear theory of permanence for deterministic systems to stochastic difference and differential equations. Our condition for coexistence requires that there is a fixed set of weights associated with the interacting populations and this weighted combination of populations' invasion rates is positive for any (ergodic) stationary distribution associated with a subcollection of populations. Here, an invasion rate corresponds to an average per-capita growth rate along a stationary distribution. When this condition holds and there is sufficient noise in the system, we show that the populations approach a unique positive stationary distribution. Moreover, we show that our coexistence criterion is robust to small perturbations of the model functions. Using this theory, we illustrate that (i) environmental noise enhances or inhibits coexistence in communities with rock-paper-scissor dynamics depending on correlations between interspecific demographic rates, (ii) stochastic variation in mortality rates has no effect on the coexistence criteria for discrete-time

S. J. Schreiber (✉)
Department of Evolution and Ecology and the Center for Population Biology,
University of California, Davis, CA 95616, USA
e-mail: sschreiber@ucdavis.edu

M. Benaïm · K. A. S. Atchadé
Institut de Mathématiques, Université de Neuchâtel, Neuchâtel, Switzerland
e-mail: michel.benaim@unine.ch

K. A. S. Atchadé
e-mail: kolawole.atchade@unine.ch

Lotka–Volterra communities, and (iii) random forcing can promote genetic diversity in the presence of exploitative interactions.

One day is fine, the next is black.—The Clash

Mathematics Subject Classification (2000)   92D25 · 60J05 · 60H10

1 Introduction

The interplay between biotic interactions and environmental fluctuations plays a crucial role in determining species richness and genetic diversity (Gillespie 1973; Chesson and Warner 1981; Turelli 1981; Chesson 1994; Ellner and Sasaki 1996; Abrams et al. 1998; Bjornstad and Grenfell 2001; Kuang and Chesson 2008, 2009). For example, competition for limited resources (Gause 1934) or sharing common predators (Holt 1977) may result in species or genotypes being displaced. However, random forcing of these systems can reverse these trends and, thereby, enhance diversity (Gillespie and Guess 1978; Chesson and Warner 1981; Abrams et al. 1998). Conversely, differential predation can mediate coexistence between competitors (Paine 1966; Holt et al. 1994; Chesson and Kuang 2008), yet environmental fluctuations can disrupt this coexistence mechanism. A fruitful approach to study this interplay is developing stochastic difference or differential equations and analyzing the long-term behavior of the probability distribution of the population sizes (Turelli 1981; Chesson 1982; Gard 1984; Chesson and Ellner 1989; Ellner 1989; Gyllenberg et al. 1994a,b; Schreiber 2007; Benaïm et al. 2008).

An intuitive approach to the problem of coexistence is given by considering the average per-capita growth rate of a population when rare (Turelli 1978; Gard 1984; Chesson and Ellner 1989). When this growth rate is positive, the population can increase and "invade" the system. For pairwise interactions, one expects that coexistence is ensured if each population can invade when it is rare and the other population is common. Indeed, Gard (1984) and Chesson and Ellner (1989) have shown for predator–prey interactions and competitive interactions that "mutual invasibility" ensures coexistence in the sense of stochastic boundedness (Chesson 1978, 1982): the long-term distribution of each population is bounded below by a positive random variable. Going beyond pairwise interactions, this mutual invasibility criterion suggests that coexistence should occur if a missing population can invade any subcommunity of the interacting populations. Surprisingly, this criterion false even for deterministic systems. May and Leonard (1975) showed with numerical simulations that competing species exhibiting a rock-paper-scissor dynamic need not coexist despite every subcommunity being invadable by a missing species.

Starting in the late 1970s, mathematicians developed a coexistence theory for deterministic models that could handle rock-paper-scissor type dynamics (Schuster et al. 1979; Hofbauer 1981; Hutson 1984; Butler and Waltman 1986; Hofbauer and So 1989; Hutson and Schmitt 1992; Jansen and Sigmund 1998; Schreiber 2000, 2006; Garay and Hofbauer 2003; Hofbauer and Schreiber 2004). Their notion of coexistence, known as permanence or uniform persistence, ensures that populations coexist despite frequent small perturbations or rare large perturbations (Jansen and Sigmund 1998; Schreiber

2006). A sufficient condition for permanence is the existence of a fixed set of weights associated with the interacting populations such that this weighted combination of populations' invasion rates is positive for any invariant measure associated with a sub-collection of populations (Hofbauer 1981; Schreiber 2000; Garay and Hofbauer 2003). Conversely, if there is a convex combination of the invasion rates that is negative for all invariant measures associated with a sub-collection of populations, then the community has one or more populations that is extinction prone (Garay and Hofbauer 2003; Hofbauer and Schreiber 2004).

While environmental stochasticity is often cited as a motivation for the concept of permanence (Hutson and Schmitt 1992; Jansen and Sigmund 1998), only recently has the effect of environmental stochasticity on permanent systems been investigated. Benaïm et al. (2008) found if a deterministic continuous-time model satisfies the aforementioned permanence criterion, then, under a suitable non-degeneracy assumption, the corresponding stochastic differential equation with a small diffusion term has a positive stationary distribution concentrated on the positive global attractor of the deterministic system. Consequently, permanent systems persist despite continual, but on average small, random perturbations. Conversely, if the deterministic dynamics satisfies the impermanence criterion, then the stochastic dynamics almost surely converges to the boundary of the state space. This asymptotic loss of one or more species occurs even if there is a positive attractor for the underlying deterministic dynamics.

For many systems, stochastic perturbations may not be small and these perturbations may not be best described by stochastic differential equations (Turelli 1978). Here, in Sects. 2 and 3, we develop a natural generalization of the permanence criteria for stochastic difference and differential equations with arbitrary levels of noise. The proofs of these results are presented in the Appendices. In Sect. 4, we develop applications of these results to competitive lottery models, discrete-time Lotka–Volterra models with environmental disturbances, and stochastic replicator equations. In particular, we show how environmental stochasticity can enhance or disrupt diversity in these models.

2 Discrete time models

2.1 The models

We study the dynamics of k interacting populations in a random environment. Let X_t^i denote the density of the i-th population at time t and $X_t = (X_t^1, \ldots, X_t^k)$ the vector of population densities at time t.[1] To account for environmental fluctuations, we introduce a random variable ξ_t that represents the state of the environment (e.g. temperature, nutrient availability) at time t. The fitness $f_i(X_t, \xi_{t+1})$ of population i at time t depends on the population state and environmental state at time $t + 1$. Under these assumptions, we arrive at the following stochastic difference equation:

$$X_{t+1} = f(X_t, \xi_{t+1}) \circ X_t \tag{1}$$

---

[1] For sequences of random vectors, we use subscripts to denote time t and superscripts to denote components of the vector. For all other vectors, we use subscripts to denote components of the vector.

where $f(x, \xi) = (f_1(x, \xi), \ldots, f_k(x, \xi))$ is the vector of fitnesses and $\circ$ denotes the Hadamard product i.e. component-wise multiplication.

Regarding (1), we make four assumptions:

A1: There exists a compact set $\mathbf{S}$ of $\mathbf{R}_+^k = \{x \in \mathbf{R}^k : x_i \geq 0\}$ such that $X_t \in \mathbf{S}$ for all $t \geq 0$.

A2: $\{\xi_t\}_{t=0}^\infty$ is a sequence of i.i.d random variables independent of X_0 taking values in a probability space E equipped with a σ-field and probability measure m.

A3: $f_i(x, \xi)$ are strictly positive functions, continuous in x and measurable in ξ.

A4: For all i, $\sup_{x \in \mathbf{S}} \int (\log f_i(x, \xi))^2 \, m(d\xi) < \infty$

Assumption **A1** ensures that the populations remain bounded for all time. Assumptions **A2** and **A3** imply that $\{X_t\}_{t=0}^\infty$ is a Markov chain on $\mathbf{S}$ and that $\{X_t\}_{t=0}^\infty$ is *Feller*, meaning that Ph, as defined below, is continuous whenever h is continuous. Assumption **A4** is a technical assumption meet by many models.

2.2 Some ergodic theory

In order to state our main results, we introduce some notation and review some basic concepts from ergodic theory. For any Borel set $A \subset \mathbf{S}$ and $x \in \mathbf{S}$, let

$$\mathbb{P}_x[X_t \in A] = \mathbb{P}[X_t \in A | X_0 = x].$$

be the probability X_t is in A given that $X_0 = x$. For various notions of convergence, it is useful to consider how the expected value of an "observable" (a function h from $\mathbf{S} \to \mathbf{R}$) depends on the dynamics of X_t. Give a bounded or nonnegative measurable function $h : \mathbf{S} \mapsto \mathbf{R}$, define

$$\mathbb{E}[h(X_1)|X_0 = x] = \int h(f(x, \xi) \circ x) m(d\xi)$$

to be the expected value of h in the next time step given that the current state of the population is x. Let P be the operator on bounded measurable functions defined by $Ph(x) = \mathbb{E}[h(X_1)|X_0 = x]$.

To understand the long-term statistical behavior of the population dynamics, it is useful to introduce invariant probability measures. Roughly, a probability measure μ is invariant if the population initially follows the distribution of μ, then it follows this distribution for all time i.e. if $\mathbb{P}[X_0 \in A] = \mu(A)$ for all Borel sets $A \subset \mathbf{S}$, then $\mathbb{P}[X_t \in A] = \mu(A)$ for all t and all Borel sets $A \subset \mathbf{S}$. One can phrase this invariance in terms of observables $h : \mathbf{S} \to \mathbf{R}$. If X_0 follows the distribution of μ, then the expected value of $h(X_0)$ equals $\int_\mathbf{S} h(x)\mu(dx)$ and the expected value of $h(X_1)$ equals $\int_\mathbf{S} Ph(x)\mu(dx)$. It follows that a Borel probability measure μ is *invariant* for $\{X_t\}_{t=0}^\infty$ or P if

$$\int_\mathbf{S} h(x)\mu(dx) = \int_\mathbf{S} Ph(x)\mu(dx) \tag{2}$$

for all continuous bounded functions $h : \mathbf{S} \to \mathbf{R}$. We let $\mathcal{P}$ denote the space of Borel probability measures on $\mathbf{S}$.

If X_t initially follows the distribution of an invariant probability measures μ, then Birkhoff's ergodic theorem implies that the temporal averages of an observable along a population trajectory converges with probability one. More precisely, if $h : \mathbf{S} \to \mathbf{R}$ is a measurable function with $\int_{\mathbf{S}} |h(x)| \mu(dx) < \infty$, then there exists a measurable function $\widetilde{h} : \mathbf{S} \to \mathbf{R}$ such that $\int_{\mathbf{S}} |\widetilde{h}(x)| \mu(dx) < \infty$ and

$$\lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t-1} h(X_s) = \widetilde{h}(X_0)$$

with probability one. When $\widetilde{h}$ is a constant function for all bounded measurable h, μ is called an *ergodic probability measure* in which case

$$\lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t-1} h(X_s) = \int_{\mathbf{S}} h(x) \mu(dx) \tag{3}$$

with probability one. Since $\int_{\mathbf{S}} h(x) \mu(dx)$ corresponds to the expected value of $h(X_0)$, Eq. (3) can be interpreted as a law of large numbers for X_t.

While the Birkhoff ergodic theorem provides a relatively complete picture of the long-term statistical behavior of X_t, it only does so when X_0 is initially distributed like an invariant probability measure. However, as we are interested in the long-term behavior of X_t for any positive initial condition X_0, new results are needed that require the concept of an invasion rate.

2.3 Invasion rates

The *expected per-capita growth rate at state x* of population i is

$$\lambda_i(x) = \int \log f_i(x, \xi) m(d\xi).$$

When $\lambda_i(x) > 0$, the i-th population tends to increase when the current population state is x. When $\lambda_i(x) < 0$, the i-th population tends to decrease when the current population state is x. For an invariant probability measure μ, we define the *invasion rate of species i with respect to μ* to be

$$\lambda_i(\mu) = \int_{\mathbf{S}} \lambda_i(x) \mu(dx)$$

The following proposition clarifies why $\lambda_i(\mu)$ is called an invasion rate. Its proof is in Appendix A.

Proposition 1 *Let μ be an invariant probability measure and $i \in \{1, \ldots, k\}$. Then there exists a bounded map $\hat{\lambda}_i : \mathbf{S} \to \mathbf{R}$ such that*:

(i) *With probability one and for μ-almost every $x \in \mathbf{S}$*

$$\lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t-1} \log f_i(X_s, \xi_{s+1}) = \hat{\lambda}_i(x) \quad \text{when } X_0 = x;$$

(ii) $\int_{\mathbf{S}} \hat{\lambda}_i(x)\mu(dx) = \lambda_i(\mu)$; *Furthermore if μ is ergodic, then $\hat{\lambda}_i(x) = \lambda_i(\mu)\mu$-almost surely.*
(iii) *If $\mu(\{x \in \mathbf{S} : x_i > 0\}) = 1$, then $\lambda_i(\mu) = 0$.*

When μ is ergodic, $\lambda_i(\mu)$ is the long-term time average of the per-capita growth rate of population i. Moreover, since each of the set $\{x_i = 0\}$ is invariant under the dynamics in (1), there exists a set $\text{supp}(\mu) \subset \{1, \ldots, k\}$ such that $x_i > 0$ if and only if $i \in \text{supp}(\mu)$ for μ-almost all x. One can interpret $\text{supp}(\mu)$ as the set of populations supported by μ. Quite intuitively, Proposition 1 implies that the long-term average of the per-capita growth is zero for all populations supported by μ i.e. $\lambda_i(\mu) = 0$ for all $i \in \text{supp}(\mu)$.

2.4 Persistence

To quantify persistence, there are two ways to think about the asymptotic behavior of $\{X_t\}_{t=0}^{\infty}$. First, one can ask what is the distribution of X_t far into the future. For example, what is the probability that the population density of each state is greater than ϵ in the long term i.e. $\mathbb{P}[X_t \geq (\epsilon, \ldots, \epsilon)]$ for large t? The answer to this question provides information what happens across many independent realizations of the population dynamics. Alternatively, one might be interested about the statistics associated with a single realization of the process i.e. a single time series. For instance, one could ask what fraction of the time was the density of each population state greater than ϵ? To answer this question, it is useful to introduce the *occupation measures*

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

where δ_{X_s} denotes a Dirac measure at X_s i.e. $\delta_{X_s}(A) = 1$ if $X_s \in A$ and 0 otherwise for any (Borel) set $A \subset \mathbf{S}$. One can interpret $\Pi_t(A)$ as the proportion of time the population spends in A up to time t.

Our first theorem addresses persistence from the second perspective. To state this theorem, for $\eta > 0$, let $\mathbf{S}_\eta = \{x \in S : x_i \leq \eta \text{ for some } i\}$ be the set of the states where at least one population has an abundance less than or equal to η. $\mathbf{S}_0$ corresponds to the states where one or more populations is extinct.

Theorem 1 (Persistence) *Assume that one of the following equivalent conditions hold*:

(i) *For all invariant probability measures μ supported on $\mathbf{S}_0$,*

$$\lambda_*(\mu) := \max_i \lambda_i(\mu) > 0$$

(ii) *There exists $p \in \Delta$ such that*

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

for all ergodic probability measures μ supported by $\mathbf{S}_0$.

Then for all $\epsilon > 0$ there exists $\eta > 0$ such that

$$\limsup_{t \to \infty} \Pi_t(\mathbf{S}_\eta) \leq \epsilon \text{ almost surely}$$

whenever $X_0 = x \in \mathbf{S}\backslash\mathbf{S}_0$.

Theorem 1 implies that fraction of time spent by the populations in $\mathbf{S}_\eta$ goes to zero as η goes to zero. Theorem 1, however, does not ensure that there is a unique positive stationary distribution. For this stronger conclusion, there has to be sufficient noise in the system to ensure after enough time any positive population state can move close to any other positive population state. More precisely, given $\eta > 0$, we say that $\{X_t\}$ is *irreducible* over $\mathbf{S}\backslash\mathbf{S}_\eta$ if there exists a probability measure Φ on $\mathbf{S}\backslash\mathbf{S}_\eta$ such that for all $x \in \mathbf{S}\backslash\mathbf{S}_\eta$ and every Borel set $A \subset \mathbf{S}\backslash\mathbf{S}_\eta$ there exists $n \geq 1$ (depending on x and A) such that

$$\mathbb{P}_x(X_n \in A) > 0$$

whenever $\Phi(A) > 0$.

Theorem 2 (Uniqueness) *Assume that $\{X_t\}$ is irreducible over $\mathbf{S}\backslash\mathbf{S}_\eta$ for all $\eta > 0$, and that the assumption of Theorem 1 holds. Then there exists a unique invariant probability measure π such that $\pi(\mathbf{S}_0) = 0$ and the occupation measures Π_t converge almost surely to π as $t \to \infty$, whenever $X_0 = x \in \mathbf{S}\backslash\mathbf{S}_0$.*

Theorem 2 ensures the asymptotic distribution of one realization of the population dynamics is given by the positive stationary distribution π. Hence, π provides information about the long-term frequencies that a population trajectory spends in any part of the population state space. To gain information about the distribution of X_t across many realizations of the population dynamics, we need a stronger irreducibility condition. This stronger condition requires that after a fixed amount of time independent of initial condition, any positive population state can move close to any other positive population state. More precisely, we say that $\{X_t\}$ is *strongly irreducible* over $\mathbf{S}\backslash\mathbf{S}_\eta$ if there exists a probability measure Φ on $\mathbf{S}\backslash\mathbf{S}_\eta$, an integer $n \geq 1$ and some number $0 < h \leq 1$ such that for all $x \in \mathbf{S}\backslash\mathbf{S}_\eta$ and every Borel set $A \subset \mathbf{S}\backslash\mathbf{S}_\eta$

$$\mathbb{P}_x(X_n \in A) \geq h\Phi(A).$$

To state the next result given $\mu, \nu \in \mathcal{P}$ define

$$\|\mu - \nu\| = \sup_B |\mu(B) - \nu(B)|$$

where the supremum is taken over all Borel sets $B \subset \mathbf{S} \backslash \mathbf{S}_0$.

Theorem 3 (Convergence in distribution) *In addition to the assumptions of Theorem 2 assume that* $\{X_t\}$ *is strongly irreducible over* $\mathbf{S} \backslash \mathbf{S}_\eta$ *for all* $\eta > 0$. *Then the distribution of* X_t *converges to* π *as* $t \to \infty$ *whenever* $X_0 = x \in \mathbf{S} \backslash \mathbf{S}_0$; *that is*

$$\lim_{t \to \infty} \|\mathbb{P}_x[X_t \in \cdot] - \pi\| = 0 \quad \text{for all } x \in \mathbf{S} \backslash \mathbf{S}_0.$$

Remark 1 Suppose that there exists a nonzero continuous function $\rho : \mathbf{S} \backslash \mathbf{S}_0 \times \mathbf{S} \backslash \mathbf{S}_0 \mapsto \mathbf{R}_+$ an integer $n \geq 1$ and a probability ν over $\mathbf{S} \backslash \mathbf{S}_0$ such that for all $x \in \mathbf{S} \backslash \mathbf{S}_0$ and Borel set $A \subset \mathbf{S} \backslash \mathbf{S}_0$

$$\mathbb{P}_x[X_n \in A] \geq \int_A \rho(x, y)\nu(dy).$$

Then $\{X_t\}$ is strongly irreducible over $\mathbf{S} \backslash \mathbf{S}_\eta$ for all $\eta > 0$.

2.5 Robust persistence

Under an additional assumption, our main condition ensuring persistence is robust to small variations of the model. The importance of this robustness stems from the fact that all models are approximations to reality. Consequently, if nearby models (e.g. more realistic models) are not persistent despite the focal model being persistent, then one can draw few (if any!) conclusions about the persistence of the modeled biological system. To state our result about robustness, let $g(x, \xi) = g_1(x, \xi), \ldots, g_k(x, \xi)$ be fitness functions. The model

$$X_{t+1} = g(X_t, \xi_{t+1}) \circ X_t \tag{4}$$

is called a δ-*perturbation of* (1) provided g satisfies conditions **A3–A4** and

$$\sup_{x \in \mathbf{S}} \mathbb{E}[\|f(x, \xi) - g(x, \xi)\|] = \sup_{x \in \mathbf{S}} \int \|f(x, \xi) - g(x, \xi)\| m(d\xi) \leq \delta.$$

Proposition 2 *Assume the dynamics* (1) *satisfies hypothesis* (i) *of Theorem* 1 *and there exist constants* $0 < \alpha \leq \beta < \infty$ *such that* $\alpha \leq f_i(x, \xi) \leq \beta$ *for all* i, x, ξ. *Then there exists* $\delta > 0$ *such that every* δ-*perturbation of* (1) *satisfies hypothesis* (i).

3 Continuous time models

For stochastic differential equation models, we assume, for presentational clarity, that $\mathbf{S} = \Delta = \{x \in \mathbf{R}_+^k : \sum_i x_i = 1\}$. However, our results hold more generally for a compact region that is forward invariant for the stochastic dynamics. The population dynamics on Δ consists of a "drift" term that describes the dynamics in the absence of noise and a "diffusion" term that describes the effects of environmental stochasticity on the population dynamics. The drift term for population i is given by $X_t^i F_i(X_t)$ where F_i is its per-capita growth rate. To allow for correlations of environmental fluctuations across populations, we assume the environmental noise is generated by an m-dimensional standard Brownian motion $(B_t^1, \ldots, B_t^m)$ and per-capita effect of B_t^j on the growth of population j is given by $\Sigma_i^j(X_t)$. Under these assumptions, we arrive stochastic differential equations of the form

$$dX_t^i = X_t^i[F_i(X_t)dt + \sum_{j=1}^m \Sigma_i^j(X_t)dB_t^j], \quad i = 1, \ldots, k. \tag{5}$$

To ensure existence and uniqueness of solutions of (5), we assume that F_i and Σ_i^j are real valued Lipschitz continuous maps on Δ. To ensure that the population dynamics remain on Δ (i.e. Δ is invariant), we assume that for each $x \in \Delta$, the drift vector $x \circ F(x)$ and the diffusion terms

$$S^j(x) = x \circ \Sigma^j(x), \quad j = 1, \ldots, m, \tag{6}$$

are elements of the tangent space $T\Delta = \{u \in \mathbf{R}^k : \sum_{j=1}^k u_j = 0\}$ of Δ.

The stochastic differential equation (5) defines a continuous time Markov process $\{X_t\}_{t\geq 0}$ on Δ. We let $\{P_t\}_{t\geq 0}$ denote the associated semigroup defined by

$$P_t h(x) = \mathbb{E}[h(X_t)|X_0 = x]$$

for every bounded or nonnegative measurable function $h : \Delta \to \mathbf{R}$. $P_t h(x)$ is the expected value of h at time t given that initial population state is x. A probability μ on Δ is called *invariant* (respectively *ergodic*) provided it is invariant (respectively ergodic) for P_t for all $t > 0$. The *occupation measure* of $\{X_t\}_{t\geq 0}$ is the measure

$$\Pi_t = \frac{1}{t} \int_0^t \delta_{X_s} ds.$$

$\Pi_t(A)$ corresponds to the fraction of time spent in the set $A \subset \Delta$ by time t.

The analog of the per-capita growth rate for these continuous time processes is given by

$$\lambda_i(x) = F_i(x) - \frac{1}{2}a_{ii}(x) \tag{7}$$

where

$$a_{ij}(x) = \sum_{k=1}^{m} \Sigma_i^k(x) \Sigma_j^k(x).$$

When $\lambda_i(x) > 0$, the population tends to increase. When $\lambda_i(x) < 0$, the population tends to decrease. Like in the discrete-time case, define

$$\lambda_i(\mu) = \int \lambda_i(x)\mu(dx) \tag{8}$$

and

$$\lambda_*(\mu) = \max_i \lambda_i(\mu). \tag{9}$$

In Appendix B, we prove the following continuous-time analog of Theorem 1

Theorem 4 *Assume that one of the equivalent conditions* (i) *and* (ii) *of Theorem* 1 *hold where* $\lambda_i(\mu)$ *and* $\lambda_*(\mu)$ *are given by formulaes* (7), (8) *and* (9). *Then the conclusion of Theorem* 1 *hold for the occupation measure of the process* $\{X_t\}_{t \geq 0}$ *solution to* (5).

To ensure the existence of a unique stationary distribution and convergence toward this distribution, we need an appropriate irreducibility condition that ensures the noise can locally push the dynamics in all directions. More precisely, we call the system (5) *nondegenerate* if the column vectors $S^1(x), \ldots, S^m(x)$ span $T\Delta$ for all $x \in \Delta \setminus \Delta_0$.

Theorem 5 *Assume that* (5) *is* non-degenerate *and the assumption of Theorem* 4 *holds. Then there exists a unique invariant probability* π *such that* $\pi(\Delta_0) = 0$. *Furthermore,*

(i) *The distribution of* X_t *converges to* π *as* $t \to \infty$ *whenever* $X_0 = x \in \Delta \setminus \Delta_0$; *that is*

$$\lim_{t \to \infty} \|\mathbb{P}_x[X_t \in \cdot] - \pi\| = 0 \ \ \text{for all } x \in \Delta \setminus \Delta_0,$$

(ii) *The occupation measures* $\Pi_t = \frac{1}{t}\int_0^t \delta_{X_s}ds$ *converge almost surely to* π, *whenever* $X_0 = x \in \Delta \setminus \Delta_0$.

3.1 Robust persistence

Let $\tilde{F}$ and $\tilde{\Sigma}$ be real valued Lipschitz continuous maps on Δ with the property that for each $x \in \Delta$, the drift vector $x \circ \tilde{F}(x)$ and the diffusion terms $\tilde{S}^j(x) = x \circ \tilde{\Sigma}^j(x)$, are elements of $T\Delta$. The stochastic differential equation

$$dX_t^i = X_t^i[\tilde{F}_i(X_t)dt + \sum_{j=1}^{m} \tilde{\Sigma}_i^j(X_t)dB_t^j], \quad i = 1, \ldots, k; \tag{10}$$

is called a δ-*perturbation* of (5) if

$$\sup_{x \in \Delta} \|F(x) - \tilde{F}(x)\| + \|\Sigma(x) - \tilde{\Sigma}(x)\| \leq \delta.$$

Proposition 3 *Assume that the dynamics* (5) *satisfies hypothesis* (*i*) *of Theorem* 1. *Then there exists* $\delta > 0$ *such that every* δ-perturbation *of* (5) *satisfies hypothesis* (*i*).

In the case $\Sigma(x) = 0$, Proposition 3 combined with Theorem 4 or 5 implies that every sufficiently small random perturbation of the deterministic system

$$\frac{dx}{dt} = x \circ F(x) \tag{11}$$

is persistent, provided that

$$\sup_{\mu} \int F_i(x)\mu(dx) > 0$$

where the supremum is taken over the invariant probabilities of (11) supported by Δ_0. This fact was also proved in Benaïm et al. (2008, Theorem 3.1) using other techniques. Hence, Theorems 4, 5 and Proposition 3 extend Benaïm et al. (2008, Theorem 3.1) beyond small perturbation of deterministic dynamics. We remark, however, the result obtained in Benaïm et al. (2008) provides an exponential rate of convergence toward π. It would be nice to see whether of not such a rate can be obtained under the more general assumptions of Theorem 5.

4 Applications

4.1 Lottery models and the storage effect

The lottery model of Chesson and Warner (1981) represents species that require a territory or "home" (an area held to the exclusion of others) in order to reproduce. Moreover, the model assumes that space is always in short supply and, consequently, all patches are occupied. Let X_t^i denote the fraction of space occupied by species i at time t. The fraction of adults of species i dying in a time step is m_i. The spaces emptied by dying individuals are immediately filled by progeny which are produced at a rate $b_i(X_t, \xi_{t+1})X_t^i$ by species i. Here ξ_t is a sequence of i.i.d. random variables that represents environmental stochasticity. If all progeny are equally likely to fill empty spaces, then the probability an empty space is filled by species i equals

$$\frac{b_i(X_t, \xi_{t+1})X_t^i}{\sum_{j=1}^{k} b_j(X_t, \xi_{t+1})X_t^j}.$$

Under these assumptions, the dynamics of the competing species are given by

$$X_{t+1}^i = (1 - m_i)X_t^i + \sum_j m_j X_t^j \frac{b_i(X_t, \xi_{t+1})X_t^i}{\sum_j b_j(X_t, \xi_{t+1})X_t^j} \quad i = 1, \dots, k \quad (12)$$

on the state space Δ. For many choices of b_i and ξ_t, (12) satisfies the irreducibility assumptions of Theorems 2 and 3. For instance, the irreducibility assumptions are satisfied if $b_i(X_t, \xi_{t+1}) = \xi_{t+1}^i$ are log-normally distributed or gamma distributed.

When there are two species, Chesson (1982) analyzed this model when $b_i(X_t, \xi_t)$ do not depend on X_t. We show how our results recover Chesson's persistence criteria. The set of ergodic invariant measures on Δ_0 are the Dirac measures, $\delta_{(0,1)}$ and $\delta_{(1,0)}$, supported on the points $(0, 1)$ and $(1, 0)$, respectively. At these ergodic measures, the invasion rates are given by

μ	$\lambda_1(\mu)$	$\lambda_2(\mu)$
$\delta_{(1,0)}$	0	$\mathbb{E}\left[\log\left(1 - m_2 + m_2 \frac{b_2((1,0),\xi)}{b_1((1,0),\xi)}\right)\right]$
$\delta_{(0,1)}$	$\mathbb{E}\left[\log\left(1 - m_1 + m_1 \frac{b_1((0,1),\xi)}{b_2((0,1),\xi)}\right)\right]$	0

Theorem 1 ensures persistence if

$$\mathbb{E}\left[\log\left(1 - m_1 + m_1 \frac{b_1((0,1),\xi)}{b_2((0,1),\xi)}\right)\right] > 0 \quad \text{and} \quad \mathbb{E}\left[\log\left(1 - m_2 + m_2 \frac{b_2((1,0),\xi)}{b_1((1,0),\xi)}\right)\right] > 0 \quad (13)$$

Hence, we have recovered the "mutual invasibility" condition for persistence of Chesson without making any monotonicity assumptions about the functions $b_i(x, \xi)$ (see also Ellner (1984)).

When the per-capita reproductive rates are frequency-independent i.e. $b_i(x, \xi_{t+1}) = \xi_{t+1}^i$ and $\xi_t = (\xi_t^1, \xi_t^2)$, the persistence condition (13) can be used to illustrate what Chesson and Warner (1981) call the "storage-effect." In the absence of environmental stochasticity, Chesson (1982) has shown that coexistence is not possible. If there is environmental stochasticity and all individuals die between generations i.e. $m_i = 1$ for all i, then (13) simplifies to

$$\mathbb{E}[\log \xi^1] > \mathbb{E}[\log \xi^2] \quad \text{and} \quad \mathbb{E}[\log \xi^2] > \mathbb{E}[\log \xi^1]$$

Both of these conditions can not be meet in which case Chesson (1982, Thm. 3.5) has shown that one of the species goes extinct with probability one. Hence, when individuals are short lived, coexistence does not occur. On the other hand, if individuals are long-lived i.e. $m_i \approx 0$ for all i, then the approximation $\log(1 + x) = x + \mathcal{O}(x^2)$ applied to (13) yields the persistence criterion

$$\mathbb{E}\left[\frac{\xi^1}{\xi^2}\right] > 1 \quad \text{and} \quad \mathbb{E}\left[\frac{\xi^2}{\xi^1}\right] > 1.$$

This condition is meet when the species exhibit some temporal partitioning: the reproductive rates ξ_t^1 and ξ_t^2 are not overly correlated. Intuitively, long-lived individuals, unlike short-lived individuals, can "store" their numbers over periods of poor conditions and, thereby, take advantage of future good conditions. This ability to store for the future in conjunction with temporal partitioning mediates coexistence.

For lottery models with three or more species, persistence criteria can be more subtle as they may require determining the invasion rates at non-trivial ergodic measures. However, an interesting exception occurs with a rock-paper-scissor version of the lottery model i.e. species B displaces species A, C displaces B, and A displaces C. To model this intransitive interaction, we assume that the per-capita reproductive rates are linear functions of the species frequencies

$$b_i(X_t, \xi_{t+1}) = \sum_j \xi_{t+1}^{ij} X_t^j$$

where

$$\xi_t = \begin{pmatrix} \beta_t & \alpha_t & \gamma_t \\ \gamma_t & \beta_t & \alpha_t \\ \alpha_t & \gamma_t & \beta_t \end{pmatrix}$$

where $\alpha_t > \beta_t > \gamma_t > 0$ for all t. For simplicity, we assume that $m^i = m$ for all i.

For any pair of strategies, say 1 and 2, the dominant strategy, 1 is this case, displaces the subordinate strategy. Indeed, assume $X_0^3 = 0$. If $y_t = X_t^2/X_t^1$ and $z_t = \sum_i \xi_{t+1}^i X_t^i$, then

$$y_{t+1} = \frac{(1-m)z_t + m(\gamma_{t+1}X_t^1 + \beta_{t+1}X_t^2)}{(1-m)z_t + m(\beta_{t+1}X_t^1 + \alpha_{t+1}X_t^2)} y_t < y_t$$

is a decreasing sequence that converges to 0. Hence, the only ergodic measures on Δ_0 are Dirac measures δ_x supported on $x = (1,0,0)$, $(0,1,0)$, $(0,0,1)$. At these ergodic measures, the invasion rates are given by

μ	$\lambda_1(\mu)$	$\lambda_2(\mu)$	$\lambda_3(\mu)$
$\delta_{(1,0,0)}$	0	$\mathbb{E}\left[\log\left(1-m+m\,\alpha_t/\beta_t\right)\right]$	$\mathbb{E}\left[\log\left(1-m+m\,\gamma_t/\beta_t\right)\right]$
$\delta_{(0,1,0)}$	$\mathbb{E}\left[\log\left(1-m+m\,\gamma_t/\beta_t\right)\right]$	0	$\mathbb{E}\left[\log\left(1-m+m\,\alpha_t/\beta_t\right)\right]$
$\delta_{(0,0,1)}$	$\mathbb{E}\left[\log\left(1-m+m\,\alpha_t/\beta_t\right)\right]$	$\mathbb{E}\left[\log\left(1-m+m\,\gamma_t/\beta_t\right)\right]$	0

A straightforward algebraic competition reveals that the conditions for persistence are satisfied if and only if

$$I(m) := \mathbb{E}\left[\log\left(1-m+m\alpha_t/\beta_t\right)\right] + \mathbb{E}\left[\log\left(1-m+m\gamma_t/\beta_t\right)\right] > 0 \quad (14)$$

We conjecture that if the opposite inequality holds, then persistence does not occur.

To see the role of the storage effect for these rock-paper-scissor communities, we can examine how the sign of $I(m)$ depends on m. Since $I(0) = 0$ and $I''(m) < 0$ for $0 \le m \le 1$, $I(m) > 0$ for a non-empty interval of m values if and only if

$$I'(0) = \mathbb{E}\left[\frac{\alpha_t}{\beta_t} + \frac{\gamma_t}{\beta_t}\right] - 2 > 0 \tag{15}$$

Moreover, if

$$I(1) = \mathbb{E}\left[\log\frac{\alpha_t}{\beta_t} + \log\frac{\gamma_t}{\beta_t}\right] > 0 \tag{16}$$

then the community persists for all $0 < m \leq 1$. However, if (15) holds but (16) does not, then the community persists for $0 < m < m^*$ for some $m^* < 1$. Hence, as in the two species example of Chesson and Warner, competitive communities with intransitives are more likely to coexist if individuals have longer generation times. However, unlike the example of Chesson and Warner, environmental noise can disrupt as well as enhance coexistence (see discussion in Sect. 5).

4.2 Discrete Lotka–Volterra dynamics with disturbances

Consider k interacting species whose dynamics in the absence of environmental disturbances is given by

$$X_{t+1} = X_t \circ \exp(AX_t + b)$$

where the matrix A describes pairwise interactions between species and b describes the intrinsic rates of growth of each species. These dynamics of this system were studied by Hofbauer et al. (1987). To account for stochastic disturbances of these dynamics, we assume that the fraction of individuals of species i surviving environmental disturbances is $\xi_t^i \in (0, 1]$ at time t. Then the dynamics become

$$X_{t+1} = X_t \circ \exp(AX_t + b) \circ \xi_{t+1} \tag{17}$$

An algebraic characterization of boundedness in terms of the matrices A and b remains an open problem even without the stochasticity. Hofbauer et al. (1987) defined the interaction matrix A to be *hierarchically ordered* if there exists a reordering of the indices such that $A_{ii} < 0$ for all i and $A_{ij} \leq 0$ whenever $i \leq j$. While this assumption excludes all types of mutualistic interactions, it allows for any type of predator–prey or competitive interaction. The following lemma extends work of Hofbauer et al. (1987) by showing that hierarchically ordered systems remain bounded in the presence of environmental disturbances. For these systems, the irreducibility conditions of Theorems 2 and 3 are meet whenever ξ_t has a positive, continuous density on the interval $(0, 1)$.

Lemma 1 *If* (17) *is hierarchically ordered, then there exists $K > 0$ such that $X_t \in [0, K]^k$ for $t \geq k + 1$.*

Proof Following Hofbauer et al. (1987) observe that

$$X_{t+1}^1 \leq X_t^1 \exp(A_{11}X_t^1 + b_1)$$

for all t as $A_{1j} \le 0$ for all $j \ge 2$ and $\xi_t^1 \le 1$. Hence, X_{t+1}^1 is bounded above by $K_1 = -\exp(b_1 - 1)/A_{11}$, and $X_t^1 \in [0, K_1]$ for $t \ge 2$.

Assume that there exist K_i such that $X_t^i \in [0, K_i]$ for $i \le j - 1$ and $t \ge i + 1$. We will show that there exists K_j such that $X_t^j \in [0, K_j]$ for $t \ge j + 1$. Indeed, by the hierarchically ordered assumption and our inductive assumption,

$$ X_{t+1}^j \le X_t^j \exp\left(A_{jj} X_t^j + b_j + \sum_{i<j} |A_{ji}| K_j \right) $$

for $t \ge j$. Hence, $X_t^j \le K_j$ for $t \ge j + 1$ where $K_j = -\exp(b_j + \sum_{i<j} |A_{ji}| K_j - 1)/A_{jj}$. Defining $K = \max K_j$ completes the proof. $\qquad\square$

The following lemma shows that verifying persistence can reduce to a linear algebra problem. In particular, this lemma implies that all the permanence criteria developed by Hofbauer et al. (1987) for hierarchal systems extends to these stochastically perturbed Lotka–Volterra systems.

Lemma 2 *Let μ be an ergodic measure for* (17) *and $I \subset \{1, \ldots, k\}$ be such that $\mu(\{x \in \mathbf{S} : x^i > 0 \text{ iff } i \in I\}) = 1$. Define $\beta_i = b_i + \mathbb{E}[\log \xi_t^i]$. If there exists a unique solution $\hat{x}$ to*

$$ \sum_j A_{ij} \hat{x}_j + \beta_i = 0 \quad \text{for } i \in I \quad \text{and} \quad \hat{x}_i = 0 \quad \text{for } i \notin I $$

then

$$ \lambda_i(\mu) = \begin{cases} 0 & \text{if } i \in I \\ \sum_j A_{ij} \hat{x}_j + \beta_i & \text{otherwise.} \end{cases} $$

Proof Let μ and I be as assumed in the statement of the lemma. Assertion (iii) of Proposition 1 implies that

$$ 0 = \lambda_i(\mu) = \sum_j A_{ij} \int x_j \mu(dx) + \beta_i $$

for all $i \in I$. Since we have assumed there is a unique solution to this system of linear equations, it follows that $\int x_i \mu(dx) = \hat{x}_i$ for all i and the lemma follows. $\qquad\square$

4.3 Stochastic replicator dynamics

A particular case of the continuous time equations (5) is given by the *stochastic replicator dynamics* introduced by Fudenberg and Harris (1992). Assume that the *fitness* of population i is described by a function $f_i : \Delta \mapsto \mathbf{R}$ of the state variable x and that

the per capita growth rate of the number of individuals in population i is stochastic, given by sum of the fitness of strategy i and a standard Brownian motion $B_i(t)$:

$$dY_t^i = Y_t^i \left(f_i(X_t) + \sigma_i dB_t^i \right), \tag{18}$$

where $X_t^i = Y_t^i / \sum_j Y_t^j$ and $\sigma_i > 0$. Then the law of motion for the state X_t can be obtained via a straightforward application of Ito's formula and takes the form (5) with

$$F_i(x) = f_i(x) - \sigma_i^2 x_i - \sum_j x_j (f_j(x) - \sigma_j^2 x_j)$$

and

$$\Sigma_i^j(x) = (\delta_{ij} - x_j)\sigma_j.$$

If there are only two types (i.e. $k = 2$), then the only ergodic measures on Δ_0 are the Dirac measures at $x = (1, 0)$ and $x = (0, 1)$, respectively. At these Dirac measures, the invasion rates are given by

μ	$\lambda_1(\mu)$	$\lambda_2(\mu)$
$\delta_{(1,0)}$	0	$f_2(1, 0) - f_1(1, 0) - \frac{1}{2}(\sigma_2^2 - \sigma_1^2)$
$\delta_{(0,1)}$	$f_1(0, 1) - f_2(0, 1) - \frac{1}{2}(\sigma_1^2 - \sigma_2^2)$	0

Hence, both strategies persist if

$$f_1(0, 1) - \frac{\sigma_1^2}{2} > f_2(0, 1) - \frac{\sigma_2^2}{2}$$

and

$$f_2(1, 0) - \frac{\sigma_2^2}{2} > f_1(1, 0) - \frac{\sigma_1^2}{2}$$

When these inequalities are satisfied, one can solve explicitly for the density of the positive stationary distribution of $X^1 = x$ on $[0, 1]$ (see, e.g., Kimura 1964)

$$\rho(x) = \frac{C}{V(x)} \exp\left(2 \int \frac{F_1(x, 1 - x, 0)}{V(x)} dx\right)$$

where $V(x) = x^2(1 - x)^2(\sigma_1^2 + \sigma_2^2)$ and C is a normalization constant. For example, if f_i are linear functions, then this stationary distribution is given by a beta distribution as we illustrate in the next example.

Since we can solve for non-trivial ergodic measures for two interacting types, we can derive explicit conditions for persistence of three interacting types. As an illustration, consider three interacting types with per-capita growth rates $f_1(x) = r_1 + b x_3$, $f_2(x) = r_2$, and $f_3(x) = r_3 - c x_1$. Here, interactions between types 1 and

3 provide a benefit $b > 0$ to type 1 and a cost $c > 0$ to type 3. To allow for coexistence, we assume the following tradeoff $r_3 - \sigma_3^2/2 > r_2 - \sigma_2^2/2 > r_1 - \sigma_1^2/2$.

Our analysis begins with pair-wise interactions. When type 1 is not present i.e. $Y_0^1 = 0$, the remaining types $i = 2, 3$ exhibit geometric Brownian motions of the form $Y_t^i = Y_0^i \exp\left((r_i - \sigma_i^2/2)t + \sigma_i B_t^i\right)$ where B_t^i are independent Brownian motions. Since $r_3 - \sigma_3^2/2 > r_2 - \sigma_2^2/2$, X_t converges almost surely to $(0, 0, 1)$ whenever $Y_0^1 = 0$ and $Y_0^3 > 0$. Similarly, when type 3 isn't present i.e. $Y_0^3 = 0$, the remaining types $i = 1, 2$ exhibit geometric Brownian motions and X_t converges almost surely to $(0, 1, 0)$ whenever $Y_0^2 > 0$. To determine the outcome of the pairwise interaction between genotypes 1 and 3, we need the invasion rates

$$\lambda_1(0, 0, 1) = r_1 + b - \sigma_1^2/2 - r_3 + \sigma_3^2/2 \quad \text{and} \quad \lambda_3(1, 0, 0) = r_3 - c - \sigma_3^2/2 - r_1 + \sigma_1^2/2$$

Both of these invasion rates are positive provided that

$$b > r_3 - \sigma_3^2/2 - r_1 + \sigma_1^2/2 > c \tag{19}$$

When (19) holds, there is a unique invariant measure μ on $\{x \in \Delta : x_1 x_3 > 0, x_2 = 0\}$ whose density $\rho(x_1) := \rho(x_1, 1 - x_1)$ is given by

$$\rho(x_1) = \frac{C}{x_1^2(1-x_1^2)(\sigma_1^2+\sigma_2^2)} \exp\left(2\int \frac{r_1 + b(1-x_1) - r_3 + cx_1 + \sigma_2^2(1-x_1) - \sigma_1^2 x_1}{x_1^2(1-x_1^2)(\sigma_1^2+\sigma_2^2)} dx\right)$$

which upon integration yields

$$\rho(x_1) = \frac{x_1^{\alpha-1} x_3^{\beta-1}}{B(\alpha, \beta)} \tag{20}$$

where $B(\alpha, \beta)$ is a normalization constant and

$$\alpha = \frac{2(\sigma_3^2 - r_3 + r_1 + b)}{\sigma_1^2 + \sigma_3^2} - 1 \quad \beta = \frac{2(\sigma_1^2 + r_3 - r_1 - c)}{\sigma_1^2 + \sigma_3^2} - 1.$$

[Fudenberg and Harris](1992, Proposition 1) provide a detailed derivation of this stationary distribution for linear f_1 and f_3.

To understand the fate of the three interacting genotypes, there are (generically) three cases to consider. First, assume that (19) is satisfied. The invasion rate for type 2 at the invariant measure μ, see (20), is given by

$$\lambda_2(\mu) = \frac{b\sigma_3^2 - (b-c)\sigma_2^2 - c\sigma_1^2 - 2br_3 + 2(b-c)r_2 + 2cr_1 + 2bc}{2(b-c)} \tag{21}$$

Whenever $\lambda_2(\mu) > 0$, Theorem 4 ensures there is a unique positive stationary distribution on Δ by choosing $p_3 \gg p_2 \gg p_1 > 0$. Since (19) implies $b - c > 0$, (21) implies that stochastic fluctuations in genotype 3's per-capita growth rate can mediate

coexistence, while stochastic fluctuations in the per-capita growth rates of the other two genotypes can disrupt coexistence.

Next, assume that (19) doesn't hold. If $b < r_3 - \sigma_3^2/2 - r_2 + \sigma_2^2/2$, then the invasion rates $\lambda_1(0, 0, 1)$ and $\lambda_2(0, 0, 1)$ are both negative and we conjecture coexistence doesn't occur. Alternatively if $b, c > r_3 - \sigma_3^2/2 - r_2 + \sigma_2^2/2$, then the boundary dynamics exhibit a rock-paper-scissor dynamic and the only ergodic invariant measures are the Dirac measures at the vertices. At these ergodic measures, the invasion rates are given by

μ	$\lambda_1(\mu)$	$\lambda_2(\mu)$	$\lambda_3(\mu)$
$\delta_{(1,0,0)}$	0	$r_2-\sigma_2^2/2-r_1+\sigma_1^2/2$	$r_3-c-\sigma_3^2/2-r_2+\sigma_2^2/2$
$\delta_{(0,1,0)}$	$r_1-\sigma_1^2/2-r_2+\sigma_2^2/2$	0	$r_3-\sigma_3^2/2-r_2+\sigma_2^2/2$
$\delta_{(0,0,1)}$	$r_1+b-\sigma_1^2/2-r_3+\sigma_3^2/2$	$r_2-\sigma_2^2/2-r_3+\sigma_3^2/2$	0

A standard computation yields that the persistence criterion is satisfied when product of the positive invasion rates is greater than the product of the absolute value of the negative invasion rates. This occurs when $b > c$. Hence, for this rock-paper-scissor dynamic, environmental stochasticity has no effect on coexistence.

5 Discussion

Understanding under what conditions interacting populations, whether they be plants, animals, or viral particles, coexist is a question of theoretical and practical importance in population biology. Both biotic interactions and environmental fluctuations are key factors that facilitate or disrupt coexistence. To better understand this interplay between these deterministic and stochastic forces, we develop a mathematical theory extending the nonlinear theory of permanence for deterministic systems to randomly forced nonlinear systems. This theory provides a biologically interpretable criterion for coexistence in the sense of stochastic boundedness (Chesson 1978, 1982). Using this theory, we illustrate that environmental noise enhances or inhibits coexistence in communities with rock-paper-scissor dynamics, has no effect on coexistence in certain Lotka–Volterra communities, and can promote or inhibit genetic diversity.

Our condition for coexistence requires that there is a fixed set of weights associated with the interacting populations and this weighted combination of populations' invasion rates is positive for any (ergodic) stationary distribution associated with a subcollection of populations. This criterion is the stochastic analog of a permanence criterion for deterministic systems (Hofbauer 1981; Schreiber 2000; Garay and Hofbauer 2003). Since these invasion rates, defined as the average per-capita growth rates on the stationary distribution, equal zero for populations supported by the stationary distribution, this criterion requires that a missing population has a positive invasion rate. Hence, for pair-wise interactions, this criterion reduces to the "mutual invasibility" criterion. When this condition holds and there is sufficient noise in the system (i.e. irreducible), we have shown the populations approach a unique positive stationary distribution whenever all types are initially present. Hence, the probability that the abundance of any population falls below a critical threshold is arbitrarily small for

sufficiently small thresholds. Moreover, the fraction of time any population spends below this threshold is arbitrarily small for sufficiently small thresholds.

The need for this generalization of the mutual invasibility criterion is illustrated in the deterministic literature by models of communities exhibiting rock-paper-scissor type dynamics (Hofbauer and Sigmund 1998). Here, we extended this analysis to stochastic counterparts of these models. If we assume that dominant strategies (e.g. rock) in these models gain a benefit b_t when playing subordinate strategies (e.g. scissor) and subordinate strategies pay a cost c_t when playing dominant strategies, then our coexistence condition (15) for long-lived individuals becomes

$$\mathbb{E}\left[\frac{b_t}{\beta_t}\right] > \mathbb{E}\left[\frac{c_t}{\beta_t}\right] \tag{22}$$

where β_t is the "base" payoff. The effect of stochasticity in b_t, c_t, and β_t on whether this criterion holds depends on the correlations between the various payoffs. Negative correlations between base payoffs and benefits (i.e. getting large benefits when base payoffs are small) makes (22) more likely to hold, while negative correlations between base payoffs and costs make it less likely to hold. Hence, stochasticity can facilitate coexistence when there are negative correlations between benefits and base-payoffs, but inhibit coexistence when there are positive correlations between benefits and base-payoffs.

For three competing genotypes in which the genotypes with the highest per-capita growth rate is exploited by the genotype with the lowest per-capita growth rate, we have shown that the effect of environmental fluctuations on coexistence is subtle. When the three genotypes exhibit a rock-paper-scissor dynamic, stochastic fluctuations have no effect on the coexistence criterion; coexistence requires that the benefit to the exploiter exceed the cost paid by the exploited. When there is no rock-paper-scissor dynamic, fluctuations in the per-capita growth rate of the exploited genotype can enhance diversity, while fluctuations in the other two genotypes can disrupt coexistence. Since this noise-induced coexistence occurs in populations with overlapping generations (i.e. a stochastic differential equation model), these results partially support Ellner and Sasaki (1996)'s assertion that "fluctuating selection can readily maintain genetic variance in species where generations overlap in such a way that only a fraction of the population is exposed to selection."

We also have shown that stochastic variation in mortality or disturbance rates have no effect the coexistence criteria for discrete-time Lotka Volterra models developed by Hofbauer et al. (1987). This surprising outcome stems from the fact that the per-capita growth rates in these models are linear functions in the population abundances. Adding non-linearities (e.g. predator saturation) to the per-capita growth rates will alter this conclusion, but the nature of this alteration remains to be understood.

Numerous mathematical challenges remain at this interface of random forcing and biotic interactions. For example, do the same criteria hold when there are temporal correlations in the environmental variables? We suspect the answer is yes. Alternatively, we conjecture that inverting the coexistence criterion (i.e. a convex combination of invasion is negative for all stationary distributions supporting subsets of species) implies an asymptotic approach to extinction with probability one. While this

conjecture has been proven for stochastic differential equations with small diffusion terms (Benaïm et al. 2008), it needs to be shown for stochastic difference equations or stochastic differential equations with large diffusion terms. Finally, only recently have invasion-based permanence criteria been developed for deterministic models of structured interacting populations (Hofbauer and Schreiber 2010). These structured models can account for heterogeneity amongst individuals in terms of the location, size, and age. Developing a mathematical framework to deal with these heterogeneities is an exciting challenge that would help us understand how interactions between individual heterogeneity, temporal heterogeneity, and biotic interactions determine diversity.

Appendix A: Proofs for discrete time models

A.1 Proof of Proposition 1

Lemma 3 *Let $g : \mathbf{S} \times E \mapsto \mathbf{R}$ be a measurable map such that $\sup_{x \in C} \int g(x, \xi)^2 m(d\xi) < \infty$. Define $\bar{g}(x) = \int g(x, \xi) m(d\xi)$. Then*

(i) *For all $x \in \mathbf{S}$ and $X_0 = x$*

$$\lim_{t \to \infty} \frac{\sum_{s=0}^{t-1} g(X_s, \xi_{s+1}) - \sum_{s=0}^{t-1} \bar{g}(X_s)}{t} = 0.$$

with probability one.

(ii) *Let μ be an invariant (respectively ergodic) probability measure for (X_t), then there exists a bounded measurable map $\hat{g}$ such that with probability one and for μ-almost every x*

$$\lim_{t \to \infty} \frac{\sum_{s=0}^{t-1} g(X_s, \xi_{s+1})}{t} = \lim_{t \to \infty} \frac{\sum_{s=0}^{t-1} \bar{g}(X_s)}{t} = \hat{g}(x) \quad \text{when } X_0 = x.$$

Furthermore

$$\int \bar{g}(x)\mu(dx) = \int \hat{g}(x)\mu(dx)$$

$$(\text{respectively } \hat{g}(x) = \int \bar{g}(x)\mu(dx) \quad \mu\text{-almost surely}).$$

Proof The first assertion follows from the strong law of large number for martingales, since $g(X_s, \xi_{s+1}) - Pg(X_s)$ is a square integrable martingale difference. The second assertion follows from Birkhoff's ergodic theorem applied to stationary Markov Chains (see Meyn and Tweedie 1993, Theorem 17.1.2). □

The first two assertions of Proposition 1 follow directly from the preceding lemma applied to $g(x, \xi) = \log f_i(x, \xi)$. For the third assertion, notice that, by assertion (i) of Proposition 1

$$\lim_{t \to \infty} \frac{\log X_t^i}{t} = \hat{\lambda}_i(x)$$

for μ-almost all $x \in \mathbf{S} \backslash \mathbf{S}_0$. Let $\mathbf{S}^{i,\eta} = \{x \in \mathbf{S} : x_i \geq \eta\}$ and $\eta^* > 0$ be such that $\mu(\mathbf{S}^{i,\eta}) > 0$ for all $\eta \leq \eta^*$. By Poincaré Recurrence Theorem, for μ almost all $x \in \mathbf{S}^{i,\eta}$

$$\mathbb{P}_x[X_t \in \mathbf{S}^{i,\eta} \text{ infinitely often}] = 1$$

for $\eta \leq \eta^*$. Thus $\hat{\lambda}_i(x) = 0$ for μ-almost all $x \in \mathbf{S}^{i,\eta}$ with $\eta \leq \eta^*$. Hence $\hat{\lambda}^i(x) = 0$ for μ-almost all $x \in \bigcup_{n \in \mathbf{N}} \mathbf{S}^{i,1/n} = \{x \in \mathbf{S} : x_i > 0\}$. This proves assertion (iii).

A.2 Proof of Theorem 1

The proof of the first assertion of the theorem follows from the following lemma.

Lemma 4 *The following two conditions are equivalent*:

(i) *For all invariant probability measures μ supported on $\mathbf{S}_0$,*

$$\lambda_*(\mu) := \max_i \lambda_i(\mu) > 0$$

(ii) *There exists $p \in \Delta$ such that*

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

for all ergodic probability measures μ supported by $\mathbf{S}_0$.

Proof To see the equivalence of the conditions we need the following version of the minimax theorem (see, e.g., Simmons 1998):

Theorem 6 (Minimax theorem) *Let A, B be Hausdorff topological vector spaces and let $\Gamma : A \times B \to \mathbf{R}$ be a continuous bilinear function. Finally, let E and F be nonempty, convex, compact subsets of A and B, respectively. Then*

$$\min_{a \in E} \max_{b \in F} \Gamma(a, b) = \max_{b \in F} \min_{a \in E} \Gamma(a, b)$$

We have that

$$\min_\mu \max_i \lambda_i(\mu) = \min_\mu \max_{p \in \Delta} \sum_i p_i \lambda_i(\mu)$$

where the minimum is taken over invariant probability measures μ with support in $\mathbf{S}_0$. Define A to be the dual space to the space bounded continuous functions from $\mathbf{S}_0$ to $\mathbf{R}$ and define $B = \mathbf{R}^k$. Let $D \subset A$ be the set of invariant probability measures and $E = \Delta$. With these choices, the Minimax theorem implies that

$$\min_{\mu} \max_{i} \lambda_i(\mu) = \max_{p \in \Delta} \min_{\mu} \sum_{i} p_i \lambda_i(\mu) \tag{23}$$

where the minimum is taken over invariant probability measures μ with support in $\mathbf{S}_0$. By the ergodic decomposition theorem Mañé (1983), the minimum of the right hand side of (23) is attained at an ergodic probability measure with support in $\mathbf{S}_0$. Thus, the equivalence of the conditions is established. $\square$

The proof of the second assertion of the theorem follows from the next two lemmas.

Lemma 5 *For all $\epsilon > 0$, there exists a $\eta > 0$ such that $\mu(\mathbf{S}_\eta) \leq \epsilon$ for every invariant probability measure μ with $\mu(\mathbf{S} \backslash \mathbf{S}_0) = 1$.*

Proof Suppose to the contrary, there exists $\epsilon > 0$ and invariant measures μ_n such that $\mu_n(\mathbf{S} \backslash \mathbf{S}_0) = 1$ and $\mu(\mathbf{S}_{1/n}) > \epsilon$ for all $n \geq 1$. By Proposition 1, $\lambda_*(\mu_n) = 0$ for all n. Let μ be a weak* limit point of these measures. Then $\mu(\mathbf{S}_0) \geq \epsilon$ and $\lambda_*(\mu) = 0$. Since $\mathbf{S} = \mathbf{S}_0 \cup \mathbf{S} \backslash \mathbf{S}_0$ and $\mathbf{S}_0, \mathbf{S} \backslash \mathbf{S}_0$ are invariant, there exists $\alpha > 0$ such that $\mu = \alpha \nu_0 + (1 - \alpha)\nu_1$ where ν_i are invariant measures satisfying $\nu_0(\mathbf{S}_0) = 1$ and $\nu_1(\mathbf{S} \backslash \mathbf{S}_0) = 1$. By Proposition 1, $\lambda_*(\nu_1) = 0$. By assumption, $\lambda_*(\nu_0) > 0$. Hence, $0 = \lambda_*(\mu) \geq \alpha \lambda_*(\nu_0) > 0$, a contradiction. $\square$

Lemma 6 *For all $x \in \mathbf{S} \backslash \mathbf{S}_0$, with probability one the set of weak* limit points of Π_t is a nonempty compact set consisting of invariant probabilities μ such that $\mu(\mathbf{S} \backslash \mathbf{S}_0) = 1$.*

Proof The process $\{X_t\}_{t=0}^{\infty}$ being a Feller Markov chain over a compact set $\mathbf{S}$, the set of weak* limit points of $\{\Pi_t\}_{t=0}^{\infty}$ is almost surely a non-empty compact subset of $\mathcal{P}$ consisting of invariant probabilities. To see why this latter point is true, let $h : \mathbf{S} \to \mathbf{R}$ be continuous function and define

$$g(x, \xi) = h(x \circ f(x, \xi)) \quad \text{and} \quad \bar{g}(x) = \int_{\mathbf{S}} g(x, \xi) m(d\xi).$$

Since $X_{t+1} = X_t \circ f(X_t, \xi_{t+1})$, $h(X_{t+1}) = g(X_t, \xi_{t+1})$ and

$$\lim_{t \to \infty} \int_{\mathbf{S}} h(x) \Pi_t(dx) - \int_{\mathbf{S}} Ph(x) \Pi_t(dx)$$

$$= \lim_{t \to \infty} \frac{1}{t} \left(\sum_{s=1}^{t} h(X_s) - \int_{\mathbf{S}} h(f(X_s, \xi) \circ X_s) m(d\xi) \right)$$

$$= \lim_{t \to \infty} \frac{1}{t} \left(\sum_{s=0}^{t-1} g(X_s, \xi_{s+1}) - \bar{g}(X_s) \right) + \frac{1}{t} (\bar{g}(X_t) - \bar{g}(X_0))$$

$$= 0 \quad \text{almost surely.}$$

where the last line follows from assertion (i) of Lemma 3. Hence, $\int_{\mathbf{S}} h(x)\mu(dx) = \int_{\mathbf{S}} Ph(x)\mu(dx)$ with probability one for weak* limit points μ of Π_t. Since $\mathbf{S}$ is compact, the set of continuous functions from $\mathbf{S}$ to $\mathbf{R}$ is separable metric space and with probability one $\int_{\mathbf{S}} h(x)\mu(dx) = \int_{\mathbf{S}} Ph(x)\mu(dx)$ for all weak* limit points of Π_t and all continuous functions $h : \mathbf{S} \to \mathbf{R}$. Thus, the weak* limit points of Π_t are almost-surely invariant probability measures.

Assertion (i) of Lemma 3 applied to $g(x, \xi) = \log(f_i(x, \xi))$ gives we have

$$\lim_{t \to \infty} \frac{\log X_t^i - \log x_i - \sum_{s=0}^{t-1} \lambda_i(X_s)}{t} = 0.$$

Since $\limsup_{t \to \infty} \frac{1}{t}\left(\log X_t^i - \log x_i\right) \leq 0$ almost surely, we get that

$$\lambda_*(\mu) \leq 0 \qquad (24)$$

almost surely for any weak* limit point μ of $\{\Pi_t\}_{t=0}^{\infty}$.

Since $\mathbf{S} = \mathbf{S}\backslash\mathbf{S}_0 \cup \mathbf{S}_0$, $\mathbf{S}\backslash\mathbf{S}_0$ is invariant, and $\mathbf{S}_0$ is invariant, there exists $\alpha \in (0, 1]$ such that $\mu = (1 - \alpha)\nu_0 + \alpha\nu_1$ where ν_0 is an invariant probability measure with $\nu_0(\mathbf{S}_0) = 1$ and ν_1 is an invariant probability measure with $\nu_1(\mathbf{S}\backslash\mathbf{S}_0) = 1$. By Proposition 1, $\lambda_i(\nu_1) = 0$ for all i. Thus, $(1 - \alpha)\lambda_i(\nu_0) \leq 0$ for all i. Since by assumption $\lambda_i(\nu_0) > 0$ for some i, α must be 1. $\qquad\square$

A.3 Proof of Theorems 2 and 3

Lemma 7 *There exists $\eta > 0$ and $\epsilon > 0$ such that*

(i) $\lambda_*(\mu) \geq \epsilon$ *for every invariant probability μ with $\mu(\mathbf{S}_\eta) = 1$, and*
(ii) $\mathbb{P}_x[X_t \notin \mathbf{S}_\eta \text{ for some } t] = 1$ *for all $x \in \mathbf{S}\backslash\mathbf{S}_0$.*

Proof To prove (i), assume to the contrary that there exists a sequence $\{\mu_n\}_{n=1}^{\infty}$ of invariant probabilities such that $\lambda_*(\mu_n) \leq 1/n$ and $\mu_n(\mathbf{S}_{1/n}) = 1$. Let μ be a weak* limit point of the $\{\mu_n\}_{n=1}^{\infty}$. Hence $\lambda_*(\mu) = 0$ by continuity of λ_*, and $\mu(\mathbf{S}_0) = 1$ since $\mu_n(\mathbf{S}_a) = 1$ for all $a > 0$ and n large enough. However, this contradicts the assumption that $\lambda_*(\mu) > 0$. Hence, there exists $\epsilon > 0$ and $\eta > 0$ such that (i) holds.

To prove (ii), let $\mathcal{E}$ be the event $\mathcal{E} = \{\forall t \geq 0 : X_t \in \mathbf{S}_\eta\}$. On $\mathcal{E}$, Π_t is almost surely supported by $\mathbf{S}_\eta$. Hence, by (i), $\lambda_*(\mu) \geq \epsilon$ almost surely on $\mathcal{E}$ for any weak* limit point μ of Π_t. This contradicts (24) in the proof of Lemma 6. Hence, $\mathcal{E}$ has probability zero, $\qquad\square$

We now pass to the proof of Theorem 2. Let $\eta > 0$ be like in Lemma 7 (ii) and Φ the probability on $\mathbf{S}\backslash\mathbf{S}_\eta$ given by the irreducibility assumption. Then, for all $x \in \mathbf{S}\backslash\mathbf{S}_0$ and every Borel set $A \subset \mathbf{S}_\eta \mathbb{P}_x[\exists n \geq 1 \ X_n \in A] > 0$ whenever $\Phi(A) > 0$. In other words, $\{X_t\}$ is a Φ-*irreducible Markov chain* on $\mathbf{S}\backslash\mathbf{S}_0$ in the sense of Meyn and Tweedie (1993, Chapter 4, Section 4.2). It then follows [see Meyn and Tweedie (1993, Proposition 10.1.1, Theorem 10.4.4)] that $\{X_t\}$ admits at most one invariant probability measure on $\mathbf{S}\backslash\mathbf{S}_0$ and Theorem 2 follows from Lemma 6.

If one now assume that $\{X_t\}$ is strongly irreducible over $\mathbf{S} \backslash \mathbf{S}_\eta$ then $\{X_t\}$ becomes Harris recurrent and aperiodic on $\mathbf{S} \backslash \mathbf{S}_0$. Since, by Theorem 2 it is a positive Harris chain, Theorem 3 follows from Orey's theorem [see Meyn and Tweedie (1993, Theorem 18.1.2)].

A.4 Proof of Proposition 2

Assume to the contrary that there exists a sequence of fitness maps $\{g_n = (g_n^1, \ldots, g_n^k)\}_{n=1}^\infty$ satisfying assumptions **A3**–**A4** such that

$$\lim_{n \to \infty} \sup_{x \in \mathbf{S}} \mathbb{E}[\|g_n(x, \xi) - f(x, \xi)\|] = 0 \qquad (25)$$

and

$$\max_i \int \log(g_n^i(x, \xi)) m(d\xi) \mu_n(dx) \leq 0$$

where μ_n is an invariant measure, supported by $\mathbf{S}_0$, for the operator P_n associated to the Markov chain

$$X_{t+1} = g_n(X_t, \xi_{t+1}) \circ X_t.$$

By compactness of $\mathbf{S}$ we may assume that $\mu_n \to \mu$ in the weak* topology. Since $\alpha \leq f_i \leq \beta$ for all i and (25) holds, it follows that

$$\lim_{n \to \infty} \int \log(g_n^i(x, \xi)) m(d\xi) \mu_n(dx) = \int \log(f_i(x, \xi)) m(d\xi) \mu(dx).$$

Hence, $\lambda_i(\mu) \leq 0$ for all i. It remains to prove that μ is invariant for P to reach a contradiction. Let $h : \mathbf{S} \mapsto \mathbf{R}$ be a continuous map. Let $\epsilon > 0$. By uniform continuity there exists $\delta > 0$ such that for all $x, u, v \in \mathbf{S}$,

$$\|u - v\| \leq \delta \Rightarrow |h(x \circ u) - h(x \circ v)| \leq \epsilon.$$

Thus

$$\begin{aligned}
|P_n h(x) - P h(x)| &= |\mathbb{E}[h(x \circ g_n(x, \xi)) - h(x \circ f(x, \xi))]| \\
&\leq 2\|h\| \mathbb{P}[\|g_n(x, \xi) - f(x, \xi)\| \geq \delta] + \epsilon \\
&\leq 2\|h\| \frac{\mathbb{E}[\|g_n(x, \xi) - f(x, \xi)\|]}{\delta} + \epsilon.
\end{aligned}$$

It then follows from (25) that $\lim_{n \to \infty} P_n h(x) = P h(x)$ uniformly in x. Therefore

$$\lim_{n \to \infty} \int P_n h(x) \mu_n(dx) = \int P h(x) \mu(dx).$$

Since by invariance of μ_n for P_n

$$\lim_{n\to\infty} \int P_n h(x)\mu_n(dx) = \lim_{n\to\infty} \int h(x)\mu_n(dx) = \lim_{n\to\infty} \int h(x)\mu(dx)$$

we get

$$\int Ph(x)\mu(dx) = \int h(x)\mu(dx),$$

proving that μ is invariant for P.

Appendix B: Proofs for the continuous time models

B.1 Proof of Theorem 4

Let $\mathcal{L}$ be the infinitesimal generator of $\{X_t\}_{t\geq 0}$. It acts on C^2 functions according to the formula

$$\mathcal{L}\psi(x) = \sum_i \frac{\partial\psi}{\partial x_i}(x)x_i F_i(x) + \mathcal{A}\psi(x) \tag{26}$$

where

$$\mathcal{A}\psi(x) = \frac{1}{2}\sum_{i,j} x_i x_j a_{ij}(x)\frac{\partial^2\psi}{\partial x_i x_j}(x) \tag{27}$$

By Ito's formulae

$$\psi(X_t) - \psi(x) - \int_0^t \mathcal{L}\psi(X_s)ds = M_t$$

is a martingale given by $M_0 = 0$ and

$$dM_t = \sum_{i=1}^k \frac{\partial}{\partial x_i}(X_t)\sum_{j=1}^m S_i^j(X_t)dB_t^j$$

where S^j is the vector given by (6). Applying this to $\psi(x) = \log(x_i)$ gives

$$\log(X_t^i) - \log(x^i) - \int_0^t \lambda_i(X_s)ds = M_t$$

with

$$dM_t = \sum_{j=1}^{m} \Sigma_i^j(X_t) dB_t^j.$$

Hence

$$d\langle M \rangle_t = \sum_{j=1}^{m} ((\Sigma_i^j(X_t))^2 dt$$

so that

$$\langle M \rangle_t \leq Ct.$$

Thus, by the strong law of large numbers for martingales,

$$\lim_{t \to \infty} \frac{\log(X_t^i) - \log(x^i) - \int_0^t \lambda_i(X_s) ds}{t} = 0$$

almost surely. The end of the proof is like the proof of Theorem 1. Details are left to the reader.

B.2 Proof of Theorem 5

By the nondegeneracy assumption, there exists (see e.g Durrett 1996, Theorem 3.8, Chapter 7) a continuous positive kernel $p_t(x, y)$ such that

$$P_t \psi(x) = \int p_t(x, y) \psi(y) dy.$$

Therefore, Theorem 1 applies to P_t for any $t > 0$.

Let π_t denote the unique positive invariant probability measure of P_t for $t > 0$. We claim that π_t is independent of t. Indeed, π_t is invariant for $P_{kt} = P_t^k$ for all $t > 0$ and $k \in \mathbb{N}$. It follows that $\pi_{k/2^n}$ is independent of k and n, and so, by the density of the dyadic rational numbers in the reals, $\pi_t = \pi$ for all $t > 0$.

Now, for any continuous bounded function ψ and any $0 \leq s < 1$,

$$|P_{n+s}\psi(x) - \pi\psi| = |P_n(P_s\psi)(x) - \pi(P_s\psi)| \leq \|P_n(x, .) - \pi\| \|P_s\psi\|_\infty$$

where $\pi\psi$ stands for $\int \psi d\pi$. Hence,

$$\lim_{n \to \infty} \|P_{n+s}(x, .) - \pi\| = 0$$

so assertion (i) of the theorem holds. The second assertion follows from the uniqueness of π and Theorem 4.

B.3 Proof of Proposition 3

Suppose (10) is a δ-perturbation of (5) with $X_0 = \tilde{X}_0 = x$. Then for all $t \geq 0$

$$
X_t - \tilde{X}_t = \int_0^t (X_s \circ F(X_s) - \tilde{X}_s \circ F(\tilde{X}_s))ds + \int_0^t (\tilde{X}_s \circ F(\tilde{X}_s) - \tilde{X}_s \circ \tilde{F}(\tilde{X}_s))ds
$$

$$
+ \int_0^t (X_s \circ \Sigma(X_s) - \tilde{X}_s \circ \Sigma(\tilde{X}_s))dB_s + \int_0^t (\tilde{X}_s \circ \Sigma(\tilde{X}_s) - \tilde{X}_s \circ \tilde{\Sigma}(\tilde{X}_s))dB_s
$$

Let

$$
v(t) = \mathbb{E}\left[\|X_t - \tilde{X}_t\|^2\right].
$$

Then, by the Cauchy-Schwartz inequality, the fact that $\|X_s\| \leq 1$, and the Ito isometry

$$
v(t) \leq 4\mathbb{E}\left[\left\|\int_0^t X_s \circ F(X_s) - \tilde{X}_s \circ F(\tilde{X}_s)ds\right\|^2 + \left\|\int_0^t \tilde{X}_s \circ F(\tilde{X}_s) - \tilde{X}_s \circ \tilde{F}(\tilde{X}_s)ds\right\|^2\right.
$$

$$
\left. + \left\|\int_0^t X_s \circ \Sigma(X_s) - \tilde{X}_s \circ \Sigma(\tilde{X}_s)dB_s\right\|^2 + \left\|\int_0^t \tilde{X}_s \circ \Sigma(\tilde{X}_s) - \tilde{X}_s \circ \tilde{\Sigma}(\tilde{X}_s)dB_s\right\|^2\right]
$$

$$
\leq 4t \int_0^t \mathbb{E}[\|X_s \circ F(X_s) - \tilde{X}_s \circ F(\tilde{X}_s))\|^2]ds + 4t \int_0^t \mathbb{E}[\|F(\tilde{X}_s) - \tilde{F}(\tilde{X}_s)\|^2]ds
$$

$$
+ 4 \int_0^t \mathbb{E}[\|X_s \circ \Sigma(X_s) - \tilde{X}_s \circ \Sigma(\tilde{X}_s)\|^2]ds + 4 \int_0^t \mathbb{E}[\|\Sigma(\tilde{X}_s) - \tilde{\Sigma}(\tilde{X}_s)\|^2]ds
$$

Using the assumption and the Lipschitz continuity of $X \circ F(X)$ and $X \circ \Sigma(X)$ it follows that, for some constant L,

$$
v(t) \leq 4tL \int_0^t v(s)ds + 4t^2\delta^2 + 4L \int_0^t v(s)ds + 4t\delta^2.
$$

Thus, for all $t \leq T$

$$
v(t) \leq A \int_0^t v(s)ds + B\delta^2
$$

where $A = 4L(T + 1)$ and $B = 4T(T + 1)$. Hence, by Gronwall's lemma

$$v(t) \leq e^{tA} B \delta^2$$

for all $t \leq T$. The remainder of proof is similar to the proof of 2. The details are left to the reader.

References

Abrams PA, Holt RD, Roth JD (1998) Apparent competition or apparent mutualism? Shared predation when populations cycle. Ecology 79(1):201–212

Benaïm M, Hofbauer J, Sandholm W (2008) Robust permanence and impermanence for the stochastic replicator dynamics. J Biol Dyn 2:180–195

Bjornstad ON, Grenfell BT (2001) Noisy clockwork: time series analysis of population fluctuations in animals. Science 293(5530):638

Butler GJ, Waltman P (1986) Persistence in dynamical systems. J Differ Equ 63:255–263

Chesson PL (1978) Predator–prey theory and variability. Annu Rev Ecol Syst 9:323–347

Chesson PL (1982) The stabilizing effect of a random environment. J Math Biol 15(1):1–36

Chesson PL (1994) Multispecies competition in variable environments. Theor Popul Biol 45(3):227–276

Chesson PL, Ellner S (1989) Invasibility and stochastic boundedness in monotonic competition models. J Math Biol 27:117–138

Chesson P, Kuang JJ (2008) The interaction between predation and competition. Nature 456(7219):235–238

Chesson PL, Warner RR (1981) Environmental variability promotes coexistence in lottery competitive systems. Am Nat 117(6):923

Durrett R (1996) Stochastic calculus. Probability and stochastics series. CRC Press, Boca Raton

Ellner SP (1984) Asymptotic behavior of some stochastic difference equation population models. J Math Biol 19:169–200

Ellner S (1989) Convergence to stationary distributions in two-species stochastic competition models. J Math Biol 27(4):451–462

Ellner S, Sasaki A (1996) Patterns of genetic polymorphism maintained by fluctuating selection with overlapping generations. Theor Popul Biol 50:31–65

Fudenberg D, Harris C (1992) Evolutionary dynamics with aggregate shocks. J Econ Theory 57:420–441

Garay BM, Hofbauer J (2003) Robust permanence for ecological differential equations, minimax, and discretizations. SIAM J Math Anal 34:1007–1039

Gard TC (1984) Persistence in stochastic food web models. Bull Math Biol 46(3):357–370

Gause GF (1934) The struggle for existence. Williams and Wilkins, Baltimore

Gillespie JH (1973) Polymorphism in random environments. Theor Popul Biol 4:193–195

Gillespie JH, Guess HA (1978) The effects of environmental autocorrelations on the progress of selection in a random environment. Am Nat 112:897–909

Gyllenberg M, Hognas G, Koski T (1994a) Null recurrence in a stochastic Ricker model. In: Analysis, algebra, and computers in mathematical research (Lulea, 1992). Lecture Notes in Pure and Applied Mathematics. Decker, New York, pp 147–164

Gyllenberg M, Hognas G, Koski T (1994b) Population models with environmental stochasticity. J Math Biol 32:93–108

Hofbauer J (1981) A general cooperation theorem for hypercycles. Monatsh Math 91:233–240

Hofbauer J, Schreiber SJ (2004) To persist or not to persist? Nonlinearity 17:1393–1406

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 repellors for maps. Proc Am Math Soc 107:1137–1142

Hofbauer J, Hutson V, Jansen W (1987) Coexistence for systems governed by difference equations of Lotka-Volterra type. J Math Biol 25(5):553–570

Holt RD (1977) Predation, apparent competition and the structure of prey communities. Theor Popul Biol 12:197–229

Holt RD, Grover J, Tilman D (1994) Simple rules for interspecific dominance in systems with exploitative and apparent competition. Am Nat 144:741–771

Hutson V (1984) A theorem on average Liapunov functions. Monatsh Math 98:267–275

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. Theor Popul Biol 54:195–201

Kimura M (1964) Diffusion models in population genetics. J Appl Probab 1:177–232

Kuang JJ, Chesson P (2008) Predation-competition interactions for seasonally recruiting species. Am Nat 171:119–133

Kuang JJ, Chesson P (2009) Coexistence of annual plants: generalist seed predation weakens the storage effect. Ecology 90:170–182

Mañé R (1983) Ergodic theory and differentiable dynamics. Springer-Verlag, New York

May RM, Leonard W (1975) Nonlinear aspects of competition between three species. SIAM J Appl Math 29:243–252

Meyn SP, Tweedie RL (1993) Markov Chains and stochastic stability. Springer, New York

Paine RT (1966) Food web complexity and species diversity. Am Nat 100:65–75

Schreiber SJ (2000) Criteria for C^r robust permanence. J Differ Equ 162:400–426

Schreiber SJ (2006) Persistence despite perturbations for interacting populations. J Theor Biol 242:844–852

Schreiber SJ (2007) On persistence and extinction of randomly perturbed dynamical systems. Discrete Contin Dyn Syst B 7:457–463

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-Verlag, Berlin

Turelli M (1978) Random environments and stochastic calculus. Theor Popul Biol 12:140–178

Turelli M (1981) Niche overlap and invasion of competitors in random environments I. Models without demographic stochasticity. Theor Popul Biol 20:1–56