Abstract. Understanding under what conditions populations, whether they be plants, animals, or viral particles, persist is an issue of theoretical and practical importance in population biology. Both biotic interactions and environmental fluctuations are key factors that can facilitate or disrupt persistence. One approach to examining the interplay between these deterministic and stochastic forces is the construction and analysis of stochastic difference equations $X_{t+1} = F(X_t, \xi_{t+1})$ where $X_t \in \mathbb{R}^k$ represents the state of the populations and $\xi_1, \xi_2, \ldots$ is a sequence of random variables representing environmental stochasticity. In the analysis of these stochastic models, many theoretical population biologists are interested in whether the models are stochastically bounded and persistent. Stochastic boundedness asserts that asymptotically $X_t$ tends to remain in compact sets. In contrast, stochastic persistence requires that $X_t$ tends to be “repelled” by some ”extinction set” $S_0 \subset \mathbb{R}^k$. Here, results on both of these properties are reviewed for single species, multiple species, and structured population models. The results are illustrated with applications to stochastic versions of the Hassell and Ricker single species models, Ricker, Beverton-Holt, and lottery models of competition, and lottery models of rock-paper-scissor games. A variety of conjectures and suggestions for future research are presented.

Accepted for publication in Journal of Difference Equations and Applications

1. Introduction

One of the most fundamental equations in population biology is “what are the necessary conditions to ensure the long-term persistence of a population or a collection of interacting populations?” A fruitful approach to addressing this question has been the development and analysis of mathematical models. For deterministic models, such as difference or differential equations, any reasonable definition of persistence requires the existence of an attractor bounded away from extinction of one or more of the populations [Schreiber, 2006]. When this attractor is a global attractor (i.e. its basin includes all non-extinction states), the system is uniformly persistent or permanent [Schuster et al., 1979, Hofbauer, 1981, Butler et al., 1986]. Permanence ensures populations recover for infrequent large perturbations often experienced by biological systems [Jansen and Sigmund, 1998, Schreiber, 2006]. Since its introduction, methods for verifying permanence have been developed extensively for deterministic models accounting for nonlinear species interactions, stage-structure within populations, and spatial heterogeneity (see, e.g, the books of Hofbauer and Sigmund [1998], Cantrell and Cosner [2003], Zhao [2003], Smith and Thieme [2011]).

Temporal fluctuations in environmental conditions also play a crucial role in determining persistence. One approach to understanding the influence of these temporal fluctuations is the study of uniform persistence for non-autonomous difference or differential equations [Thieme, 2000, Mierczyski et al., 2004, Smith and Thieme, 2011]. A strength of this approach is that it allows for relatively arbitrary forms of temporal fluctuations including periodic, quasi-periodic and stochastic motions. However, the definition of persistence in these studies often requires that the population trajectories remain uniformly bounded away from the extinction state. For many stochastic models where the vagaries of the environment are encapsulated in randomly...
varying parameters, this requirement is too strong [Gillespie, 1973, Chesson and Warner, 1981, 
Chesson, 2008, 2009]. For these models, population trajectories often drift arbitrarily close to the 
extinction set. However, under certain conditions, there may be a probabilistic tendency to stay 
away from this extinction set. When this occurs, the model is stochastically persistent [Chesson, 
1978, 1982].

Here, I provide a partial review of the latter approach to persistence in fluctuating environ-
ments. The main class of stochastic difference equations under consideration are introduced 
in Section 2. Definitions of stochastic boundedness and stochastic persistence are given. For 
the former, a rather general theorem is presented. For the latter, results are more dependent 
on model structure. Consequently, Sections 3, 4, and 5 discuss results for scalar models, mul-
tispecies models, and structured single species models, respectively. Section 6 concludes with 
parting comments and suggestions for future research.

2. Background

To study population dynamics in a random environment, consider stochastic difference equa-
tions of the form

\[ X_{t+1} = F(X_t, \xi_{t+1}) \]

where \( X_t \in S \) represents the “state” of the population at time \( t \) (e.g. a vector of densities or 
 frequencies) and \( \xi_t \) is a random variable that determines the “environmental conditions” at time 
\( t \). Throughout this article, I make the following standing assumptions.

A1: \( \{\xi_t\}_{t=0}^\infty \) is a sequence of i.i.d random variables taking values in a separable metric space 
\( \Omega \) (such as \( \mathbb{R}^n \)).

A2: \( F : S \times \Omega \to S \) is a continuous function where \( S \) is a closed subset of \( \mathbb{R}^k \).

A3: There is a closed subset \( S_0 \subset S \) such that \( S_0 \) and \( S \setminus S_0 \) are invariant i.e. \( F(x, \omega) \in S_0 \) 
if and only if \( x \in S_0 \) and \( F(x, \omega) \in S \setminus S_0 \) if and only if \( x \in S \setminus S_0 \).

Assumptions A1–A2 imply that \( \{X_t\}_{t=0}^\infty \) is a Markov chain on \( S \) and that \( \{X_t\}_{t=0}^\infty \) is Feller.

For the many of the results presented here, \( S \) is either the non-negative orthant \( \mathbb{R}_+^k \) of \( \mathbb{R}^k \) in 
which case \( x \in S \) is a vector of population densities or \( S \) is the probability simplex \( \Delta = \{x \in 
\mathbb{R}_+^k : \sum_i x_i = 1 \} \) in which case \( x \in S \) is a vector of population frequencies. \( S_0 \) in assumption 
A3 is interpreted as the “extinction set” where one or more populations have gone extinct. The 
invariance of \( S_0 \) implies that once the population has gone extinct it remains extinct i.e. the 
“no cats, no kittens” principle in population biology.

It is natural to study the asymptotic behavior of (1) from two perspectives. First, one might 
ask “what is the probability the populations are in a particular configuration far into the future?” 
More precisely, given a Borel set \( B \subset S \), what can we say about \( \mathbb{P}[X_t \in B] \) (i.e. the probability 
\( X_t \in B \) for large \( t \)? Since these probabilities correspond the frequency of observing a particular 
event across many realizations of the stochastic process, answering this question provides an 
“ensemble view” of the long-term dynamics. An alternative perspective corresponds to asking 
“for the typical population trajectory how frequently does it visit a particular configuration far 
into the future?” To answer this question, it useful to introduce the empirical measures for the 
Markov chain \( X_t \) given by

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

where \( \delta_x \) denotes a Dirac measure at the point \( x \) i.e. \( \delta_x(A) = 1 \) if \( x \in A \) and 0 otherwise. For any 
Borel set \( A \subset S \), \( \Pi_t(A) \) is the fraction of time that \( X_s \) spends in \( A \) for \( 1 \leq s \leq t \). Provided the
This review focuses on two aspects of the asymptotic behavior: stochastic boundedness in which the populations tend to stay bounded and stochastic persistence in which the populations tend to stay away from the extinction set $S_0$. Both aspects are viewed from the “ensemble” and “typical trajectory” perspectives.

2.1. Stochastic boundedness. When studying models of population dynamics, the first question that comes to mind is “are the long term population abundances bounded?” After all, we live in a finite world so population numbers can not become arbitrary large for indefinitely long periods of time. For deterministic models, an appropriate notion of boundedness is dissipativeness: the existence of a compact set $C \subset S$ such that all solutions of (1) eventually enter and remain in $C$ for all future time. While trivially met when $S$ is compact, this notion of boundedness is, in general, too strong for stochastic models with non-compact $S$. For example, theoretical population biologists often use models of the form $X_{t+1} = \xi_{t+1} X_t f(X_t)$ where $f$ is a positive function representing the survivorship and $\xi_t$ is a log-normally distributed random variable representing the mean number of offspring produced by an individual at time $t$. Since log-normal random variables are absolutely continuous on $(0, \infty)$, $X_t$ can become arbitrarily large with positive probability at any time step. While one might argue that this behavior is biologically unrealistic, models of this variety have provided many important biological insights and therefore deserve a careful mathematical treatment.

One less restrictive notion of boundedness, suggested by Chesson [1984], is that the probability of $X_t$ getting arbitrarily large becomes vanishingly small:

**Definition 1.** (1) is stochastically bounded in probability if for all $\epsilon > 0$ there exists a compact set $C \subset S$ such that

$$\liminf_{t \to \infty} \mathbb{P}[X_t \in C | X_0 = x] \geq 1 - \epsilon$$

for all $x \in S$.

This definition of boundedness implies that across many realizations of the dynamics of (1), there is a small probability that populations lie outside a compact set far into the future.

Alternatively, the empirical measure point of view insists that the fraction of time the populations spend at arbitrarily high densities becomes vanishingly small.

**Definition 2.** (1) is stochastically bounded almost surely if for all $\epsilon > 0$ there exists a compact set $C \subset S$ such that

$$\liminf_{t \to \infty} \Pi_t [X_t \in C] \geq 1 - \epsilon$$

almost surely whenever $X_0 = x \in S$.

As in the case of deterministic models, a practical method for verifying both forms of stochastic boundedness is finding an appropriate Lyapunov-type function. Recall a function $V : S \to \mathbb{R}_+$ is called proper if $\lim_{||x|| \to \infty} V(x) = +\infty$. The following theorem shows that stochastic boundedness follows if there is a proper function $V$ decreasing, on average, along population trajectories whenever population densities are high.

**Theorem 2.1.** Let $V : S \to \mathbb{R}_+$ be a continuous, proper function. If there exist Borel functions $\alpha, \beta : \Omega \to \mathbb{R}_+$ such that

$$V(F(x, \omega)) \leq \alpha(\omega) V(x) + \beta(\omega)$$

for all $\omega, x$, $\mathbb{E}[\log \alpha(\xi)] < 0$, $\mathbb{E}[\log^+ \alpha(\xi)] < \infty$, and $\mathbb{E}[\log^+ \beta(\xi)] < \infty$ where $\log^+(z) = \max\{\log(z), 0\}$, then (1) is stochastically bounded in probability and almost surely.
Remark 1. An alternative proof for the case of almost-sure stochastic boundedness was given by Benaïm and Schreiber [2009, Proposition 4].

Proof. Define
\[ Y_t = V(X_t), \, \alpha_t = \alpha(\xi_t), \text{ and } \beta_t = \beta(\xi_t). \]
Define \( Z_t \) iteratively by \( Z_0 = Y_0 \) and
\[ Z_{t+1} = \alpha_{t+1}Z_t + \beta_{t+1} \]

Theorem 2.1 in [Diaconis and Freedman, 1999] implies there exists a non-negative random variable \( \widehat{Z} \) such that \( Z_t \) converges in probability to \( \widehat{Z} \) and the empirical measures \( \frac{1}{t} \sum_{s=1}^{t} Z_s \) converge almost surely to the distribution of \( \widehat{Z} \).

Equation (4) and the definition of \( Z_t \) implies that \( Z_t \geq Y_t \geq 0 \) for all \( t \). Given \( \epsilon > 0 \), choose \( a > 0 \) such that \( \mathbb{P}[\widehat{Z} \in [0, a]] \geq 1 - \epsilon/2 \). Since \( Z_t \) converges in probability to \( \widehat{Z} \), there exists \( T > 0 \) such that
\[ \mathbb{P}[X_t \in V^{-1}([0, a])] = \mathbb{P}[Y_t \in [0, a]] \geq \mathbb{P}[Z_t \in [0, a]] \geq 1 - \epsilon \]
for all \( t \geq T \). Since \( V \) is proper, \( V^{-1}([0, a]) \) is compact. Therefore, \( X_t \) is stochastically bounded in probability.

Similarly, with probability one,
\[ \lim_{t \to \infty} \inf \mathbb{E}_t (V^{-1}([0, a])) = \lim_{t \to \infty} \inf \frac{1}{t} \sum_{s=1}^{t} \delta_{Z_s}([0, a]) \geq \lim_{t \to \infty} \inf \frac{1}{t} \sum_{s=1}^{t} \delta_{Z_s}([0, a]) = \mathbb{P}[\widehat{Z} \in [0, a]] \geq 1 - \epsilon/2 \]

Therefore \( X_t \) is stochastically bounded almost surely.

Stochastic boundedness ensures the existence of an invariant probability measure: a Borel probability measure \( \mu \) on \( S \) such that if \( X_0 \) is distributed according to \( \mu \) (i.e. \( \mathbb{P}[X_0 \in A] = \mu(A) \) for all Borel sets \( A \subset S \)), then it is distributed according to \( \mu \) for all time i.e. \( \mathbb{P}[X_n \in A] = \mu(A) \) for all Borel \( A \subset S \). The proof of the following proposition follows from a standard argument, see e.g. Duflo [1997, Proposition 6.1.8]. One can think of it as the stochastic analog of the fact that the \( \omega \)-limit set for a point is non-empty for dissipative maps.

Proposition 2.2. If (1) is bounded in probability, then the set of weak* limit points of \( \frac{1}{t} \sum_{s=0}^{t-1} \mathbb{P}[X_s \in \cdot | X_0 = x] \) with \( x \in S \) is non-empty and each of these limit points is an invariant probability measure. Alternatively, if (1) is bounded almost surely, then the set of weak* limit points of \( \Pi_t \) with \( X_0 = x \in S \) is almost-surely non-empty and each of these limit points is an invariant probability measure.

2.2. Stochastic persistence. When the population dynamics are stochastically bounded, population biologists are often interested understanding the conditions ensuring the long-term persistence of the populations. A natural analog of uniform persistence or permanence for stochastic models is given below. For these definitions, it useful to introduce the set of the population states within \( \eta > 0 \) of extinction
\[ S_{\eta} = \{ x \in S : d(x, S_0) \leq \eta \}. \]
where \( d(x, S_0) = \min_{y \in S_0} \| x - y \| \).

From the “ensemble” point of view, the following notion of stochastic persistence was introduced by Chesson [1982].
Definition 3. (1) is persistent in probability if for all $\epsilon > 0$ there exists $\eta > 0$ such that
\begin{equation}
\limsup_{t \to \infty} P[X_t \in S_\eta | X_0 = x] \leq \epsilon
\end{equation}
for all $x \in S \setminus S_0$.

This definition asserts that reaching low densities or frequencies is very unlikely far into the future. The next definition provides the “typical trajectory” perspective on persistence.

Definition 4. (1) is persistent almost surely if for all $\epsilon > 0$ there exists $\eta > 0$ such that
\begin{equation}
\limsup_{t \to \infty} \Pi_t[S_\eta] \leq \epsilon \text{ almost surely}
\end{equation}
whenever $X_0 = x \in S \setminus S_0$.

This definition asserts that the fraction of time a typical population trajectory spends near extinction states is very small.

When (1) is stochastically bounded, Proposition 2.2 implies that there exists an invariant probability measure. If in addition (1) is stochastically persistent, then the following proposition implies that there exists a positive invariant probability measure $\mu$ i.e. an invariant probability measure satisfying $\mu(S_0) = 0$.

Proposition 2.3. If (1) is stochastically persistent and bounded in probability, then the set of weak* limit points of $\frac{1}{t} \sum_{s=0}^{t-1} P[X_s \in \cdot | X_0 = x]$ with $x \in S \setminus S_0$ is non-empty and each of these limit points is a positive, invariant measure. Alternatively, if (1) is stochastically persistent and bounded almost surely, then the set of weak* limit points of $\Pi_t$ with $X_0 = x \in S \setminus S_0$ is almost-surely non-empty and each of these limit points is almost-surely a positive, invariant measure.

Proof. Suppose that (1) is persistent and bounded in probability. Let $x \in S \setminus S_0$ and assume that $t_k \uparrow \infty$ is such that $\frac{1}{t_k} \sum_{s=0}^{t_k-1} P[X_s \in \cdot | X_0 = x]$ converges in the weak* topology to $\mu$. Proposition 2.2 implies that $\mu$ is invariant. On the other hand, given any natural number $n$, stochastic persistence in probability implies there exists $\eta_n > 0$ and $T > 0$ such that
\[ P[X_t \in S_{\eta_n} | X_0 = x] < 1/n \]
for $t \geq T$. Hence, by weak* convergence $\mu(S_{\eta_n}) \leq 1/n$ and $\mu(S_0) \leq \limsup_{n \to \infty} \mu(S_{\eta_n}) = 0$.

The proof for the case of almost sure stochastic persistence and boundedness is similar.

When a unique positive invariant probability measure exists and the system is stochastically persistent, one can often show that if $X_0 = x \in S \setminus S_0$, then the distribution of $X_t$ converges to $\mu$ and $\Pi_t$ converges almost surely to $\mu$. A powerful tool for verifying this stronger form of persistence is the following theorem due to Meyn and Tweedie [2009, Chapter 15]. This theorem relies on the concept of $\varphi$-irreducibility with respect to a Borel set $B \subset S$: there exists a Borel measure $\varphi$ on $B$ such that $\varphi(A) > 0$ implies that $P[X_n \in A$ for some $n | X_0 = x] > 0$ for all $x \in B$.

Theorem 2.4. If (1) is $\varphi$-irreducible on $S \setminus S_0$ and there exists a positive function $V : S \setminus S_0 \to \mathbb{R}_+$, a compact set $C \subset S \setminus S_0$, and constant $\beta > 0$ such that
\[ E[V(X_1) | X_0 = x] \leq (1 - \beta)V(x) + 1_C(x) \text{ for all } x \in S \setminus S_0, \]
then there exists a unique invariant probability measure $\mu$ and the distribution of $X_t$ converges in the weak* topology to $\mu$ and $\Pi_t$ almost surely converges in the weak* topology to $\mu$ whenever $X_0 = x \in S \setminus S_0$. Here $1_C$ is the indicator function for $C$ i.e. $1_C(x) = 1$ if $x \in C$ and 0 otherwise.
A drawback of requiring $\varphi$-reducibility is that it can be difficult to verify or demonstrably false for important classes of tractable models. For instance, many important biological insights have been gleaned from models where there are a finite number of environmental states (i.e. $\Omega$ is a finite set). These models rarely satisfy the irreducibility condition and, consequently, may not have a unique positive invariant measure.

In the next three sections, we review results for stochastic persistence of scalar single species models, multiple species models, and structured species models (e.g. spatial, age, or size structure).

3. Scalar models

The simplest form of (1) are the scalar models describing the dynamics of an unstructured, closed population i.e. $k = 1$, $S = [0, \infty)$, $S_0 = \{0\}$, in which case (1) simplifies to

\[
X_{t+1} = f(X_t, \xi_{t+1})X_t \text{ with } X_t \in [0, \infty)
\]

where $f(x, \omega) : [0, \infty) \times \Omega \to (0, \infty)$ is a continuous function. Random difference equations of this form have been studied extensively by many authors [Athreya and Dai, 2000, 2002, Athreya and Schuh, 2003, Bezandry et al., 2008, Bhattacharya and Majumdar, 2004, Blé et al., 2007, Chesson, 1982, Ellner, 1984, Fagerholm and Högnäs, 2002, Gyllenberg et al., 1994a, Haskell and Sacker, 2005, Vellekoop and Högnäs, 1997]. Here, I focus on results that relate to stochastic persistence and boundedness.

A reasonably general criterion extinction, persistence, and population blow up is given by the following Theorem. The proof of extinction and blow-up follows from standard arguments that have been used by many authors [Chesson, 1982, Ellner, 1984, Fagerholm and Högnäs, 2002, Gyllenberg et al., 1994a, Vellekoop and Högnäs, 1997]. The argument for stochastic boundedness follows from Theorem 2.1 and stochastic persistence follows from the univariate version of Theorem 1 in [Schreiber et al., in press].

**Theorem 3.1.** Assume $f(x, \omega)$ is a positive decreasing function in $x$ and $\mathbb{E}[\log^+ f(0, \xi_t)] < \infty$. Then

1. if $\mathbb{E}[\log f(0, \xi_t)] < 0$, then $\lim_{t \to \infty} X_t = 0$ for all $x \geq 0$,
2. if $\lim_{x \to \infty} \mathbb{E}[\log f(x, \xi_t)] > 0$, then $\lim_{t \to \infty} X_t = \infty$ for all $x > 0$, and
3. if $\mathbb{E}[\log f(0, \xi)] > 0$ and $\lim_{x \to \infty} \mathbb{E}[\log f(x, \xi_t)] < 0$, then (7) is stochastically bounded in probability and almost surely, and stochastically persistent almost surely.

The assumption that $f(x, \omega)$ is decreasing with $x$ holds for many “classical” single species models. However, this assumption doesn’t hold for species exhibiting an Allee effect [Courchamp et al., 1999]. Ellner [1984] proved results for models where $x \mapsto f(x, \omega)$ is not monotonic and $x \mapsto xf(x, \omega)$ is increasing. More generally, it seems reasonable to make the following conjecture.

**Conjecture 1.** If $f(x, \omega)$ is a positive function and $\mathbb{E}[\log^+ f(0, \xi_t)] < \infty$, then assertion (i) of Theorem 3.1 holds provided (7) is $\varphi$-irreducible on $(0, \infty)$ and stochastically bounded almost surely, and assertion (iii) of Theorem 3.1 holds. Moreover, if (7) is stochastically bounded in probability (or almost surely) and $\mathbb{E}[\log f(0, \xi_t)] < 0$, then (7) is not stochastically persistence in probability (or almost surely).
Proof. Assume that $\mathbb{E}[\log f(0, \xi)] < 0$ and $X_0 > 0$. Then by the Strong Law of Large Numbers,

$$\limsup_{t \to \infty} \frac{1}{t} \log X_t = \limsup_{t \to \infty} \frac{1}{t} \left( \sum_{s=0}^{t-1} \log f(X_s, \xi_{s+1}) + \log X_0 \right) \leq \lim_{t \to \infty} \frac{1}{t} \left( \sum_{s=0}^{t-1} \log f(0, \xi_{s+1}) + \log X_0 \right) = \mathbb{E}[\log f(0, \xi)] < 0$$

with probability one. Hence, $\lim_{t \to \infty} X_t = 0$ with probability one.

Assume that $\lim_{x \to \infty} \mathbb{E}[\log f(x, \xi)] > 0$ and $X_0 > 0$. Then by the Strong Law of Large Numbers,

$$\liminf_{t \to \infty} \frac{1}{t} \log X_t = \liminf_{t \to \infty} \frac{1}{t} \left( \sum_{s=0}^{t-1} \log f(X_s, \xi_{s+1}) + \log X_0 \right) \geq \lim_{t \to \infty} \frac{1}{t} \left( \sum_{s=0}^{t-1} \lim_{x \to \infty} \log f(x, \xi_{s+1}) + \log X_0 \right) = \lim_{x \to \infty} \mathbb{E}[\log f(x, \xi)] > 0$$

with probability one. Hence, $\lim_{t \to \infty} X_t = \infty$ with probability one.

Assume that $\mathbb{E}[\log f(0, \xi)] > 0$ and $\lim_{x \to \infty} \mathbb{E}[\log f(x, \xi)] < 0$. To verify boundedness, choose $M > 0$ and $\epsilon > 0$ such that $\mathbb{E}[\log f(x, \xi)] \leq -\epsilon$ for all $x \geq M$. Define $V : [0, \infty) \to [0, \infty)$ to be the identity function $V(x) = x$. Define $\alpha_t = f(M, \xi_{t+1})$. Define $\beta_t = f(0, \xi_t)M$. Since $f$ is decreasing in $M$, $X_{t+1} \leq \alpha_{t+1}X_t$ whenever $X_t \geq M$. On the other hand, $X_{t+1} \leq f(0, \xi_{t+1})M = \beta_{t+1}$ whenever $X_t \leq M$. Since $\mathbb{E}[\log + \beta_t] < \infty$ by assumption and $\mathbb{E}[\log \alpha_t] \leq -\epsilon$ by definition, Theorem 2.1 implies that $X_t$ is stochastically bounded in probability and almost surely. The proof of almost-sure persistence follows verbatim as in the proof of [Schreiber et al., in press, Theorem 1]. The compactness assumption in this proof is only needed for $S_0$ which equals $\{0\}$ for (7). \qed

Theorem 3.1 suggests the following conjecture.

**Conjecture 2.** Under the same assumptions of Theorem 3.1(iii), (7) is persistent in probability.

**Example 3.2** (The stochastic Hasell model). To illustrate the utility of Theorem 3.1, consider Hassell [1975]’s single species model in which $\omega = (\lambda, b)$ and

$$f(x, \omega) = \frac{\lambda}{1 + x^b}$$

where $\lambda$ is the intrinsic fitness of an individual, and $b > 0$ determines the strength of intraspecific competition. For this model,

$$\mathbb{E}[\log f(x, \xi)] = \mathbb{E}[\log \lambda_t] - \mathbb{E}[b_t] \log(1 + x)$$

Hence, $\lim_{x \to \infty} \mathbb{E}[\log f(x, \xi)] = -\infty$ and $\mathbb{E}[\log f(0, \xi)] = \mathbb{E}[\log \lambda_t]$. Theorem 3.1 implies almost sure extinction if $\mathbb{E}[\log \lambda_t] < 0$ and almost sure persistence if $\mathbb{E}[\log \lambda_t] > 0$. \bbox

To arrive at stronger conclusions, more assumptions about either the form of the nonlinearity or the noise are necessary. With respect to the form of the nonlinearity, Ellner [1984, Theorem 2.2] showed that a monotonicity assumption on $xf(x)$ ensures convergence in distribution to a positive random variable. In the case of the stochastic Hassell model described in Example 3.2, this convergence occurs whenever $b_t \in (0, 1]$ and $\mathbb{E}[\log \lambda_t] > 0$. 

**Theorem 3.3** (Ellner 1984). Assume that $F(x, \omega) = xf(x, \omega)$ is continuously differentiable and strictly increasing in $x$, $f(x, \omega)$ is strictly decreasing in $x$, and $\mathbb{E}[\log f(x, \xi_t)] < \infty$ for some $x > 0$. If $\mathbb{E}[\log f(0, \xi_t)] > 0$ and $\lim_{x \to \infty} \mathbb{E}[\log f(x, \xi_t)] < 0$, then there exists a positive invariant probability measure $\mu$ and the distribution of $X_t$ converges in the weak* topology to $\mu$ whenever $X_0 = x > 0$.

The assumption that $x \mapsto F(x, \omega)$ is strictly increasing implies $F(\cdot, \omega)$ is a monotone map for each $\omega$. For deterministic systems, this monotonicity provides a lot of leverage to understand the map dynamics even in higher dimensions, as reviewed in Hirsch and Smith [2005]. This leverage has been extended to random maps as reviewed by Chueshov [2002]. In the special cases of the stochastic Beverton-Holt model $X_{t+1} = \frac{\lambda_{t+1}}{1+a_{t+1}X_t}X_t$ and the stochastic Beverton-Holt model with survivorship $X_{t+1} = \frac{\lambda_{t+1}}{1+a_{t+1}X_t}X_t + s_{t+1}X_t$, a similar result to Theorem 3.3 was proven by Haskell and Sacker [2005] and Bezaudry et al. [2008], respectively.

When $b_t > 1$ for the stochastic Hassell model, monotonicity fails (i.e. the map $x \mapsto x f(x, \omega)$ is unimodal) and other assumptions are necessary to ensure convergence to a positive random variable. Vellekoop and Högnäs [1997] proved an ergodic form of stochastic persistence by placing stronger assumptions on the random variables $\xi_t$. Their result is applicable to the Hassell model under the assumption that $\lambda_t$ is constant. The proof uses the Lyapunov function characterization of ergodicity described in Theorem 2.4.

**Theorem 3.4** (Vellekoop and Högnäs 1997). Assume that $f(x, \omega) = \lambda g(x)^{-\omega}$ where $g$ is a positive differentiable function satisfying $x \mapsto x g'(x)/g(x)$ is strictly increasing on $[0, \infty)$. Assume that $\xi_t$ are i.i.d. positive random variables with $\mathbb{E}[\xi_t], \mathbb{E}[\xi_t^2] < \infty$ and a positive density on $(0, \infty)$. If $\lambda > 1$, then there exists a positive invariant probability measure $\mu$, the distribution of $X_t$ converges to $\mu$ whenever $X_0 = x > 0$, and the empirical measures $\Pi_t$ converge almost surely to $\mu$ whenever $X_0 = x > 0$.

Since $g(x) = 1 + x$ satisfies $x g'(x)/g(x) = x/(1 + x)$, Theorem 3.4 applies to the stochastic Hassell model with $\lambda_t$ constant and $\xi_t = b_t$. This theorem is also applicable to the stochastic Ricker equation $X_{t+1} = X_t \exp(r - a_{t+1}X_t)$ where $r > 0$ is the intrinsic per-capita growth rate of the population and $\xi_t = a_t$ measures the intensity of introspective competition. Gyllenberg et al. [1994] studied the stochastic Ricker model when either $r_t$ or $a_t$ vary randomly. More recently, Fagerholm and Högnäs [2002] studied the dynamics of the stochastic Ricker model when both $r_t$ and $a_t$ vary randomly. Quite surprising, they prove that if $\mathbb{E}[r_t] = 0$, then (7) is null recurrent: there exists no positive invariant probability measure despite $\mathbb{P}[X_t \in A \text{ infinitely often}]$ for all Borel sets $A$ with positive Lebesgue measure.

**Theorem 3.5** (Fagerholm & Högnäs 2002). Consider the stochastic Ricker $X_{t+1} = X_t \exp(r_{t+1} - a_{t+1}X_t)$ where

- $r_1, r_2, \ldots$ is a sequence of positive i.i.d. random variables such that $\mathbb{E}[r_1] < \infty$ and $r_t$ has a positive density on $(0, \infty)$,
- $a_1, a_2, \ldots$ is a sequence of positive i.i.d. random variables independent of $r_t$ such that $\mathbb{E}[a_t] < \infty$, and
- there exists $x_c > 0$ such that $\mathbb{E}[\exp(r_1 x)] < \infty$ for all $0 \leq x \leq x_c$.

Then one of the following hold

- **extinction**: If $\mathbb{E}[r_t] < 0$, then $X_t$ converges to 0 with probability one,
- **null recurrence**: If $\mathbb{E}[r_t] = 0$ and $X_0 > 0$, then $X_t$ is null recurrent, or
- **persistence**: if $\mathbb{E}[r_t] > 0$, then there exists a positive invariant measure $\mu$ such that the distribution of $X_t$ converges to $\mu$ in the weak* topology.
Fagerholm and Högnäs [2002] and Gyllenberg et al. [1994] also studied the case when $a_t \leq 0$ with positive probability. In this case the dynamics have the potential to be explosive (i.e. the intraspecific interactions enhance growth) and, consequently, the analysis is more subtle. This case was motivated by numerical studies where $r_t$ and $K_t$ were normally distributed.

4. Multiple species interactions

For multiple species interactions in fluctuating environments, one can consider models of the form

$$X_{t+1}^i = f_i(X_t, \xi_t)X_t^i \text{ with } i = 1, \ldots, k$$

where $X_t^i$ is the density or frequency of species $i$ at time $t$. For these models $S$ is a closed subset of $\mathbb{R}^k_+ = \{x \in \mathbb{R}^k : x_i \geq 0\}$ and $S_0 = \{x \in S : \prod x_i = 0\}$ corresponds to the extinction of one or more species. These models have been used extensively to understand under what conditions environmental stochasticity and species interactions facilitate or disrupt species or genetic diversity [Chesson, 1978, Chesson and Warner, 1981, Chesson, 1982, 1994, Kuang and Chesson, 2009, Anderies and Beisner, 2000, Turelli, 1978, 1981].

Despite extensive numerical and theoretical work on these types of models, there are (to the best of my knowledge) only two sets of mathematical results concerning stochastic persistence for these multispecies models. The first set of results [Chesson and Ellner, 1989, Ellner, 1989] applies to two species competitive systems. The second set of results [Schreiber et al., in press] apply to $k$ species systems provided the dynamics satisfy an appropriate compactness assumption. Both of these results utilize the average per-capita growth rates of populations when rare. More specifically, define the \textit{per-capita growth rate of species} $i$ \textit{at population state} $x$ to be

$$\lambda_i(x) = \mathbb{E}[\log f_i(x, \xi_t)],$$

and define the \textit{per-capita growth rate of species} $i$ \textit{at invariant probability measure} $\mu$ to be

$$\lambda_i(\mu) = \int \lambda_i(x)\mu(dx).$$

When $\mu$ is ergodic (i.e. $\int \mathbb{E}[h(F(x, \xi_t))]\mu(dx) = \int h(x)\mu(dx)$ if and only if $h$ is a constant function $\mu$-almost surely), $\lambda_i(\mu)$ is the long-term average of the per-capita growth rate of population $i$ when invading a resident system supported by the invariant measure $\mu$. More precisely, since each of the sets $\{x_i = 0\}$ is invariant under the dynamics in (8), there exists a set $\text{supp}(\mu) \subset \{1, \ldots, k\}$ such that for $\mu$-almost all $x, x_i > 0$ if and only if $i \in \text{supp}(\mu)$. One can interpret $\text{supp}(\mu)$ as the set of populations supported by $\mu$.

4.1. Competition between two species. For competitive interactions between two species, Chesson and Ellner [1989], Ellner [1989] proved a “mutual invisibility” condition implies stochastic persistence. In all of their results, they built on the single species results by assuming

- \textbf{B1:} for each $i = 1, 2$, there exists a positive invariant measure $\mu_i$ such that the distribution of $X_t^i$ converges to $\mu_i$ in the weak* topology whenever $X_0^j > 0$ and $X_0^j = 0$ for $j \neq i$, and

- \textbf{B2:} the per-capita growth rates $\lambda_i(x)$ are continuous functions. \footnote{This assumption is made in Ellner [1989] and is met for most models. A slighter weaker assumption is made in Chesson and Ellner [1989].}

The mutual invisibility condition [Turelli, 1981] asserts that the species coexist provided that $\lambda_1(\mu_2) > 0$ and $\lambda_2(\mu_1) > 0$. Intuitively, whenever one species, say species 2, is rare, the dynamics of the other species approaches its invariant measure $\mu_1$. At this invariant measure, the per-capita growth rate of species 2 is positive (i.e. $\lambda_2(\mu_1) > 0$) and, consequently, increases in abundance. Since each species increases in abundance when rare, they coexist. Under the
assumption that the competitive dynamics are monotonic, Chesson and Ellner [1989] proved
that mutual invasibility implies stochastic persistence in probability.

**Theorem 4.1** (Chesson and Ellner 1989). Assume (8) with \( k = 2 \) satisfy B1-B2, and

- the equations \( x_1 = F_1((x_1, 0), \xi_t) \) and \( x_2 = F_2((0, x_2), \xi_t) \) can be true with probability one
- only for \( x_1 = 0 \) and \( x_2 = 0 \), respectively, and
- the functions \( F_i(x, \omega) \) are non-decreasing in \( x_1 \) and \( x_2 \), and positive whenever \( x_i > 0 \).

If \( \lambda_1(\mu_2) > 0 \) and \( \lambda_2(\mu_1) > 0 \), then (8) is stochastically persistent in probability.

It is natural to make the following conjecture. This conjecture follows from Theorem 5.2
whenever the dynamics of (8) asymptotically enters a compact set.

**Conjecture 3.** Under the conditions of Theorem 4.1, (8) is almost surely persistent.

Under a stronger assumption about the noise terms \( \xi_t \) in Theorem 4.1, Ellner [1989] proved
that there exists a unique positive invariant measure \( \mu \) such that the distribution of \( X_t \) converges
to \( \mu \) whenever \( X_0^1 > 0 \) and \( X_0^2 > 0 \).

**Example 4.2.** Chesson and Ellner [1989] illustrated the applicability of Theorem 4.1 with the
following competition model

\[
X_{t+1}^1 = \frac{\zeta_{t+1}^1 X_t^1}{1 + \zeta_{t+1}^1 X_t^1 + \zeta_{t+1}^2 X_t^2} + a X_t^1
\]

\[
X_{t+1}^2 = \frac{\zeta_{t+1}^2 X_t^2}{1 + \zeta_{t+1}^1 X_t^1 + \zeta_{t+1}^2 X_t^2} + a X_t^2
\]

where \( \zeta_i^j > 0 \) represents the per-capita fecundity of species \( i \) and \( 0 < a < 1 \) represents the fraction
of individuals surviving to the next time step. \( \zeta_i^1, \zeta_i^2 \) are assumed to have an interchangeable
joint distribution (i.e. \( \mathbb{P}((\zeta_i^1, \zeta_i^2) \in A) = \mathbb{P}((\zeta_i^2, \zeta_i^1) \in A) \) for any Borel set \( A \subset \mathbb{R}_+^2 \)). If
\( \mathbb{E}[\log(\zeta_i^1 + a)] > 0 \), then Theorem 3.3 implies that B1 is satisfied. By the interchangeability
assumption \( \lambda_1(\mu_2) = \lambda_2(\mu_1) \). Following Chesson [1988], define

\[
g(\omega) = \log \left( \frac{\omega_1}{1 + \omega_2 x_2} + a \right).
\]

Chesson [1988] proved that

\[
\lambda_1(\mu_2) = -\frac{1}{2} \mathbb{E} \left[ \int_{\zeta_1^2}^{\zeta_2^2} \int_{\zeta_1^1}^{\zeta_2^1} \frac{\partial^2 g(\omega)}{\partial \omega_1 \partial \omega_2} d\omega_1 d\omega_2 \right].
\]

Since \( \frac{\partial^2 g(\omega)}{\partial \omega_1 \partial \omega_2} = -\frac{1}{\omega_1} \), it follows that \( \lambda_1(\mu_2) = \lambda_2(\mu_1) > 0 \). □

The monotonicity assumption is too strong to cover all models of competitive interactions.
For instance, Theorem 4.1 does not apply to stochastic Ricker models of competition (see Ex-
ample 4.4 below). However, Ellner [1989] proved that under a stronger assumption on the noise
terms \( \xi_t \), it is possible to show stochastic persistence for such models. The theory presented in
the next section provides a similar approach to verifying persistence for these models.

**Theorem 4.3** (Ellner 1989). Assume B1-B2 and

- (8) is \( \varphi \) irreducible on \( (0, \infty) \times (0, \infty) \),
- (8) is strongly continuous i.e. for any measurable \( A \subset \mathbb{R}_+^2 \), \( \mathbb{P}[X_1 \in A | X_0 = x_n] \) converges
to \( \mathbb{P}[X_1 \in A | X_0 = x] \) whenever \( x_n \to x \), and
- for any \( x \in \mathbb{R}_+^2 \), \( \sup_{n \geq 0} \mathbb{E}[\log^+ X_t^i | X_0 = x] < \infty \) for \( i = 1, 2 \).

If \( \lambda_1(\mu_2) > 0 \) and \( \lambda_2(\mu_1) > 0 \), then there exists a unique positive invariant measure \( \mu \) and the
distribution of \( X_t \) converges to \( \mu \) in the weak* topology whenever \( X_0^1 > 0 \) and \( X_0^2 > 0 \).
Example 4.4 (Ricker equations of competition). To illustrate the applicability of Theorem 4.3 consider a stochastic version of the Ricker equations of competition:

\begin{equation}
X_{t+1}^i = X_t^i \exp(\xi_{t+1}^i - \alpha_j X_t^j) \quad i, j = 1, 2; i \neq j
\end{equation}

where \(\alpha_j > 0\) are inter-specific competition coefficients and \(\xi_{t+1}^i\) are normally distributed intrinsic rates of growth with means \(r_i\). Since \(\exp(\xi_{t}^i)\) are log-normally distributed, they have a positive density on \((0, \infty)\). Theorem 3.5 implies that \(B1\) holds whenever \(r_i > 0\) for \(i = 1, 2\). In particular, both species are stochastically persistent in isolation under this assumption.

The positive density of \((\xi_{t+1}^1, \xi_{t+1}^2)\) on \((0, \infty) \times (0, \infty)\) implies that \((11)\) is \(\varphi\)-irreducible on \((0, \infty) \times (0, \infty)\) with respect to Lebesgue measure. The strong-continuity condition is easily verified. The boundedness condition follows from Theorem 2.1 with stochastic persistence for the competing species whenever

\[ \lambda_i(\mu_j) = r_i - \alpha_j E[\hat{X}_j] > 0 \quad \text{for} \quad i = 1, 2 \quad \text{and} \quad j \neq i \]

where \(\hat{X}_j\) is distributed like \(\mu_j\). By invariance of \(\mu_j\),

\[ E[\log \hat{X}_j] = E[\log X_0^j | X_0^j = \hat{X}_j] = E[\log(\hat{X}_0^j \exp(\xi_0^j - \hat{X}_j))] = E[\log \hat{X}_j] + r_j - E[\hat{X}^j]. \]

Hence, \(E[\hat{X}^j] = r_j\) and coexistence occurs if

\[ r_i > \alpha_j r_j \quad \text{for} \quad i = 1, 2 \quad \text{and} \quad j \neq i. \]

Thus, the conditions for coexistence are the same for this stochastic version of the Ricker equations of competition and their deterministic counterpart. ■

4.2. General multispecies. Recently, Schreiber et al. [in press] extended the results similar to Chesson and Ellner to an arbitrary number of species by developing stochastic analogs of the classical permanence criteria for deterministic systems [Hofbauer, 1981, Hofbauer and Sigmund, 1998]. These results are based on the following assumptions about \((8)\).

\begin{itemize}
  \item \textbf{C1}: There exists a compact set \(S\) of \(R^k_+ = \{x \in R^k : x_i \geq 0\}\) such that \(X_t \in S\) for all \(t \geq 0\).
  \item \textbf{C2}: \(f_i(x, \omega)\) are strictly positive functions, continuous in \(x\) and measurable in \(\omega\).
  \item \textbf{C3}: For all \(i\), \(\sup_{x \in S} E[(\log f_i(x, \xi_t))^2] < \infty\).
\end{itemize}

Assumption \textbf{C1} requires that the populations remain bounded for all time. Assumption \textbf{C2} implies that \(\{X_t\}_{t=0}^\infty\) is \textit{Feller}. Assumption \textbf{C3} is a technical assumption meet by many models.

Under these assumptions, Schreiber et al. [in press] proved that if a species can invade (i.e. has a positive average per-capita growth rate) every invariant measure supported on \(S_0\), then the system is stochastically persistent.

\textbf{Theorem 4.5} (Schreiber, Benaïm and Atchade 2011). Assume \textbf{C1-C3} and one of the following equivalent conditions hold:

\begin{itemize}
  \item[(i)] For all invariant probability measures \(\mu\) supported on \(S_0\),
  \[ \lambda_*(\mu) := \max_i \lambda_i(\mu) > 0, \]
  \item[(ii)] There exists \(p \in S\) such that
  \[ \sum_i p_i \lambda_i(\mu) > 0 \]
\end{itemize}

for all ergodic probability measures \(\mu\) supported by \(S_0\).

Then \((8)\) is almost-surely stochastically persistent.
While this theorem, as shown below, applies to many models, it has several limitations. First, it doesn’t provide a statement about persistence in probability. None the less, it is natural to make the following conjecture.

**Conjecture 4.** Under the assumptions of Theorem 4.5, (8) is persistent in probability.

A second limitation of Theorem 4.5 is that requires a dynamics are asymptotically confined to a compact set. While this limitation as discussed earlier might be biologically realistic, it would be useful to have a result that applies to stochastically bounded systems. In particular, one could ask whether the following conjecture (or an appropriate modification of it) is true.

**Conjecture 5.** Assume (8) is stochastically bounded in probability (respectively, almost surely), C1-C3 hold, and \( \lambda(\mu) > 0 \) for all invariant measures \( \mu \) supported by \( S_0 \). Then (8) is stochastically persistent in probability (respectively, almost surely).

Theorem 4.5 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.

**Theorem 4.6** (Schreiber, Benaïm and Atchade 2011). Assume that \( \{X_t\} \) is \( \varphi \)-irreducible over \( S \setminus S_\eta \) for all \( \eta > 0 \), and that the assumption of Theorem 4.5 holds. Then there exists a unique invariant probability measure \( \mu \) such that \( \mu(S_0) = 0 \) and the occupation measures \( \Pi_t \) converge almost surely to \( \mu \) as \( t \to \infty \), whenever \( X_0 = x \in S \setminus S_0 \).

Note that \( \varphi \)-irreducibility condition does not require that the same \( \varphi \) is used for all \( \eta > 0 \). Under a stronger irreducibility condition, Schreiber et al. [in press] proves that the distribution of \( X_t \) converges to \( \mu \) whenever \( X_0 \in S \setminus S_0 \).

**Example 4.7** (Coexistence of many competitors). To illustrate the applicability of Theorem 4.6 to higher-dimensional models of competition, consider the lottery model of Chesson and Warner [1981]. This model describes species requiring a territory or “home” (an area held to the exclusion of others) in order to reproduce. Moreover, space is always in short supply and, consequently, all patches are occupied. Let \( X_i^t \) denote the fraction of space occupied by species \( i \) at time \( t \), \( \xi_t^i \) the fecundity of species \( i \) at time \( t \), and \( d \) the fraction of individuals dying each time step. Under these assumptions, the lottery model is given by

\[
X_{i+1}^t = (1 - d)X_i^t + d \frac{X_i^t\xi_{i+1}^t}{\sum_j X_j^t\xi_{j+1}^t} \quad i = 1, \ldots, k.
\]

Here \( S \) is the probability simplex \( \{x \in \mathbb{R}_+^k : \sum_i x_i = 1\} \). Let the \( \log \xi_t^i \) be normally distributed with means \( b_i > 0 \) and variances \( \sigma_i^2 \geq 0 \). Furthermore, assume that \( \xi_1^1, \ldots, \xi_k^k \) are independent.

Since log-normal distributions have a positive density on \((0, \infty)\), \( \{X_t\} \) is \( \varphi \)-irreducible on \( S_\eta \) with respect to Lebesgue measure for all \( \eta > 0 \).

If \( b_1 > \cdots > b_k \) take on distinct values and there is no environmental noise (i.e. \( \sigma_i = 0 \)), then species 1 excludes all the remaining species i.e. \( \lim_{t \to \infty} X_i^t = 0 \) for \( i = 2, \ldots, k \) whenever \( X_0^1 > 0 \). To show how environmental stochasticity can alter this ecological outcome, consider the case that \( b_1 = \cdots = b_k \) and \( \sigma_i > 0 \) for all \( i \). Let \( \mu \) be any invariant probability measure on \( S_0 \). Choose a species \( i \) such that \( \mu(\{x \in S : x_j > 0 \text{ iff } j = i\}) < 1 \). By Taylor’s theorem,

\[
\lambda_i(\mu) = -d + d \int \mathbb{E} \left[ \frac{\xi_i^t}{\sum_j x_j^t\xi_j^t} \right] \mu(dx) + O(d^2).
\]
Independence and Jensen’s inequality imply

\begin{equation}
\int \mathbb{E} \left[ \frac{\xi^i_t}{\sum_j x_j \xi^j_t} \right] \mu(dx) > \mathbb{E}[\xi^i_t] \int \frac{1}{\sum_j x_j \mathbb{E}[\xi^j_t]} \mu(dx) = 1
\end{equation}

where the final equality follows from the assumption that \( b_1 = \cdots = b_k \). Combining equations (14) and (15) imply that

\( \lambda_i(\mu) > 0 \)

provided that \( d > 0 \) is sufficiently small. By compactness of the invariant probability measures on \( S_0 \), it follows that \( \lambda^*(\mu) > 0 \) for all invariant probability measures supported by \( S_0 \) whenever \( d > 0 \) is sufficiently small. Hence, Theorem 4.6 implies there is a unique positive stationary measure and (13) is persistent almost-surely. By continuity, these conclusions still apply whenever \( \max_{i,j} |b_i - b_j| \) is sufficiently small. Therefore, environmental stochasticity can mediate coexistence between an arbitrary number of competitors.

**Example 4.8** (Rock, paper, scissors). The basic lottery model described in Example 4.7 per-capita reproductive rates are independent of species frequencies. Frequency-dependent feedbacks, however, can be quite important. To illustrate these feedbacks and to illustrate that persistence may require more than invasibility by a missing species, consider a rock-paper-scissor version of the lottery model. To model this intransitive interaction, we assume that the per-capita reproductive rates are linear functions of the species frequencies

\[ b_t(X_t, \xi_{t+1}) = \sum_j \xi^i_{t+1} X^j_t \]

where

\[ \xi_t = \left( \begin{array}{ccc} \beta_t & \alpha_t & \gamma_t \\ \gamma_t & \beta_t & \alpha_t \\ \alpha_t & \gamma_t & \beta_t \end{array} \right) \]

and \( \alpha_t > \beta_t > \gamma_t > 0 \) for all \( t \). The frequency-dependent lottery model becomes

\begin{equation}
X^i_{t+1} = (1 - d)X^i_t + d \frac{X^i_t b_t(\xi_{t+1}, X_t)}{\sum_j X^j_t b_t(\xi_{t+1}, X_t)}, \quad i = 1, 2, 3.
\end{equation}

For any pair of strategies, say 1 and 2, the dominant strategy, 1 is this case, displaces the subordinate strategy. Indeed, assume \( X^3_t = 0 \). If \( y_t = X^2_t / X^1_t \) and \( z_t = \sum_i \xi^i_{t+1} X^i_t \), then

\[ y_{t+1} = \frac{(1 - d)z_t + d(\gamma_{t+1}X^1_t + \beta_{t+1}X^2_t)}{(1 - d)z_t + d(\beta_{t+1}X^1_t + \alpha_{t+1}X^2_t)} y_t < y_t \]

is a decreasing sequence that converges to 0. Hence, the only ergodic, invariant probability measures on \( \Delta_0 \) are Dirac measures \( \delta_x \) supported on \( x = (1, 0, 0), (0, 1, 0), (0, 0, 1) \). At these ergodic measures, the invasion rates are given by

\begin{align*}
\mu & \quad \lambda_1(\mu) & \lambda_2(\mu) & \lambda_3(\mu) \\
\delta_{(1,0,0)} & 0 & \mathbb{E} \left[ \log (1 - d + d \alpha_t / \beta_t) \right] & \mathbb{E} \left[ \log (1 - d + d \gamma_t / \beta_t) \right] \\
\delta_{(0,1,0)} & \mathbb{E} \left[ \log (1 - d + d \gamma_t / \beta_t) \right] & 0 & \mathbb{E} \left[ \log (1 - d + d \alpha_t / \beta_t) \right] \\
\delta_{(0,0,1)} & \mathbb{E} \left[ \log (1 - d + d \alpha_t / \beta_t) \right] & \mathbb{E} \left[ \log (1 - d + d \gamma_t / \beta_t) \right] & 0
\end{align*}

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

\begin{equation}
\mathbb{E} \left[ \log (1 - d + d \alpha_t / \beta_t) \right] + \mathbb{E} \left[ \log (1 - d + d \gamma_t / \beta_t) \right] > 0
\end{equation}
For small $d > 0$, a Taylor’s approximation similar to Example 4.7 yields the following simpler condition for persistence:

$$E \left[ \frac{\alpha_t}{\beta_t} \right] + E \left[ \frac{\gamma_t}{\beta_t} \right] > 2.$$ 

for $d > 0$ sufficiently small. ■

I conjecture that if the opposite inequality of (17) holds, then persistence does not occur.

More generally,

**Conjecture 6.** Assume that C1-C3 hold, (8) is $\varphi$-irreducible over $S \setminus S_\eta$ for all $\eta > 0$, and $\lambda^*(\mu) < 0$ for all invariant measures $\mu$ supported on $S_0$. Then

$$\lim_{t \to \infty} \prod_i X_i^t = 0$$

with probability one.

Bena¨ım et al. [2008] proved the continuous-time version of this conjecture for stochastic ODEs on the probability simplex with a small diffusion term. Without the $\varphi$-irreducibility assumption, this conjecture is definitively false.

5. **Structured populations**

Populations often consist of a heterogeneous mixture of individuals in different states such as its age, size, physiological condition, or location in space [Caswell, 2001]. If the population consists of $k$ states, then its dynamics on $S = \mathbb{R}_+^k$ can be described by nonlinear, stochastic matrix models of the form

$$(18) \quad X_{t+1} = A(X_t, \xi_{t+1})X_t$$

where $X_t = (X_1^t, \ldots, X_k^t)$ is the vector of population densities across the different states and $A(x, \omega)$ is a non-negative matrix whose entries represent transition probabilities, survivorship likelihoods, and fecundities. Here, the extinction set is $S_0 = \{0\}$.

When population abundances are low, it seems reasonable to approximate the dynamics of (18) with the linear equation

$$X_{t+1} = A(0, \xi_{t+1})X_t$$

in which case,

$$X_t = A(0, \xi_t) \cdots A(0, \xi_1)X_0.$$ 

Under suitable conditions (e.g. $A(0, \xi_t)$ are primitive, and $E(\ln ||A(0, \xi_t)||) < \infty$), the work of Ruelle [1979] and Arnold et al. [1994] implies that there exists a quantity $\gamma$ such that

$$\lim_{t \to \infty} \frac{1}{t} \ln \left( X_1^t + \cdots + X_k^t \right) = \gamma$$

with probability one.

In other words, the total population size $X_1^t + \cdots + X_k^t$ grows approximately like $(X_0^1 + \cdots + X_0^k)e^{\gamma t}$.

The quantity $\gamma$ is known as the dominant Lyapunov exponent and is also known as the stochastic growth rate in theoretical ecology [Tuljapurkar, 1990, Caswell, 2001]. For the linearized model, if $\gamma > 0$, the population grows exponentially and persists. Alternatively, if $\gamma < 0$, the population is driven to extinction.

To contend with the nonlinearities in structured population models, Hardin et al. [1988] extended the work of Ellner [1984] to structured populations in serially uncorrelated environments. Under a slightly different set of assumptions, Benaim and Schreiber [2009] proved a similar result that also addresses convergence of the empirical measures and also applies to asymptotically stationary environments. Since these latter assumptions are slightly easier to present, I will focus...
on them. In addition to standing assumptions A1-A3, the following additional assumptions are needed.

**D1: Primitivity:** There is a positive integer $T$ such that with probability one

$$A(X_t, \xi_t) \ldots A(X_1, \xi_1)$$

has all positive entries whenever $X_1 \in \mathbb{R}^k_+$ and $t \geq T$.

**D2: Smoothness:** The map $(x, \omega) \mapsto A(x, \omega)$ is Borel, and

$$x \mapsto A(x, \omega) = F(x, \omega)$$

is twice continuous differentiable for all $\omega \in \mathcal{S}$, $x \in \mathbb{R}^k_+$, and

$$\mathbb{E} \left( \sup_{\|x\| \leq 1} \ln^+ \left( \|F(x, \omega)\| + \|D_x F(x, \omega)\| + \|D^2_x F(x, \omega)\| \right) \right) < +\infty$$

**D3: Intraspecific competition:** The matrix entries $A_{ij}(x, \omega)$ satisfy

$$\frac{\partial A_{ij}}{\partial x_l}(x, \omega) \leq 0$$

for all $\omega$ and $x$. Moreover, for each $i$ there exists some $j$ and $l$ such that this inequality is strict for all $\omega$ and $x$.

**D4: Compensating density dependence:** All entries of the derivative of $x \mapsto A(x, \omega)x$ are non-negative for all $\omega$ and $x$.

Assumption D1, roughly translated, asserts that after sufficiently many time steps, individuals in every state contribute to the abundance of individuals in all other states. For constant environments (i.e. $\mathcal{S}$ consists of a single environmental state), this assumption corresponds to a matrix being primitive [Horn and Johnson, 1990, Caswell, 2001]. Assumption D2 is purely technical, but is met for most models. It is worth noting that Hardin et al. [1988] do not require smoothness. Assumption D3 accounts for competition between the stages. Assumption D4 ensures that (18) is a random, monotone dynamical system [Chueshov, 2002].

When $\gamma$ is negative for the linearization, extinction is expected as the following proposition demonstrates.

**Proposition 5.1.** Assume D1 and D3 hold. Then

$$\limsup_{n \to \infty} \frac{1}{n} \ln \|X_t\| \leq \gamma$$

with probability one. In particular, if $\gamma < 0$, then

$$\lim_{n \to \infty} X_t = 0$$

with probability one.

When $\gamma$ is positive for the linearization, stochastic persistence is expected as the following theorem shows.

**Theorem 5.2.** Assume D1-D4 and the assumptions of Theorem 2.1 hold (i.e. the system is stochastically bounded). If $\gamma > 0$, then there exists an invariant probability measure $\mu$ such that $\mu(S_0) = 0$, the distribution of $X_t$ converges to $\mu$ whenever $X_0 = x > 0$. Moreover, with probability one, $\Pi_t$ converges toward $\mu$ whenever $X_0 = x > 0$.

The monotonicity assumptions D3-D4 are definitely not necessary for stochastic persistence. Hence, it is natural to conjecture that

**Conjecture 7.** Assume D1-D2 hold and the assumptions of Theorem 2.1 hold. If $\gamma > 0$, then (18) is stochastically persistent almost surely and in probability.
To illustrate the applicability of Theorem 5.2, we consider an example from Benaïm and Schreiber [2009].

**Example 5.3 (Biennial plants).** Biennial plants typically flower only in the second year of their existence after which they die. However, for many biennial species, individual plants may exhibit delayed flowering in which they flower in a later year. Delayed flowering can serve as a bet-hedging strategy in uncertain environments [Roerdink, 1987] provided a detailed analysis of a density-independent model of delayed flowering. Here, I describe a density-dependent version of his model. Let \( X_t^1 \) denote the abundance of one year old individuals in year \( t \) and \( X_t^2 \) denote the abundance of individuals greater than one year old in year \( t \). Let \( p \) be the probability that a plant flowers during its second year. Let \( f s_1(x_1 + x_2) \) with \( s_1(x_1, x_2) = \frac{1}{1 + b_2(x_1 + x_2)} \) be the mean number of germinating seeds produced by a flowering plant. Let \( s_2(x_1, x_2) = \frac{a}{1 + b_2(x_1 + x_2)} \) be the probability that a plant survives to the next year. If we assume that \( f = \xi_t \) at time \( t \), then the plant dynamics are given by

\[
X_{t+1} = \left( \frac{0}{s_2(X_t)} \right) \frac{p \xi_{t+1} s_1(X_t)}{(1 - p) s_2(X_t)} X_t
\]

Let \( \xi_1, \xi_2, \ldots \) be a sequence of independent random variables that are gamma distributed with probability density function

\[
g(t) = \frac{1}{\theta^\theta \Gamma(k)} e^{-t/\theta},
\]

scale parameter \( \theta > 0 \), shape parameter \( k > 0 \) and \( \Gamma(k) = \int_0^\infty e^{-t} dt \). The mean and variance of \( \xi_t \) are given by \( k \theta \) and \( k \theta^2 \). Roerdink [1987] found an explicit formula for the dominant Lyapunov exponent. For \( 0 < p < 1 \), the dominant Lyapunov exponent is given by

\[
\gamma = \ln a(1 - p) + K^{-1} \int_0^\infty \ln(1 + t) t^{-k-1} e^{-zt} dt
\]

where \( K = \int_0^\infty t^{-k-1} e^{-zt} dt \) and \( z = (1 - p)^2 / (\theta p) \). For \( p = 0 \), \( \gamma = \ln a \), while for \( p = 1 \), \( \gamma = \frac{1}{2} (\ln(a\theta) + \psi(a)) \) where \( \psi(a) \) is the digamma function. Roerdink proved that \( \frac{\partial \gamma}{\partial p} \) is positive at \( p = 0 \) and approaches \( -\infty \) as \( p \) approaches 1. Hence, in the words of Roerdink, “this remarkable result implies that, independent of the ... means and variances of [the fertilities],” the stochastic growth rate of the population is maximized by the population playing an appropriate bet hedging strategy for flowering (i.e. \( p \) strictly between 0 and 1). Theorem 5.2 implies that persistence is more likely for populations playing a bet hedging strategy (i.e. \( \gamma \) is maximized at an intermediate \( p \)).

6. Concluding remarks and future directions

While these results represent the promising beginnings of a general theory for persistence of stochastic difference equations, there is still much work to be done. The conjectures sprinkled throughout this review are merely the tip of the iceberg. To give a sense of some issues that lie beneath surface, I discuss three directions for future research.

First and foremost, one can ask “Is there a general theorem unifying all the particular cases of the stochastic persistence results?” For dissipative deterministic models, there are two characterizations of uniform persistence or permanence. The average Lyapunov characterization due to Hutson [1984, 1988] requires the existence of a non-negative function that increases on average for trajectories near the extinction set. Alternatively, Butler and Waltman [1986], Garay [1989] and Hofbauer and So [1989] provided topological characterizations of uniform persistence in terms of Morse decompositions and stable sets. Intuition suggests there should be a stochastic analog of the average Lyapunov function characterization. Indeed Benaïm et al. [2008] used
average Lyapunov functions for the deterministic models to prove stochastic persistence for the stochastically perturbed models. Whether this argument can be extended is an exciting and challenging open problem.

Many environmental signals are positively autocorrelated in time [Vasseur and Yodzis, 2004]. Understanding the impacts of these autocorrelations for population persistence and species interactions is a very active area of research in theoretical and empirical population biology [Heino et al., 2000, Gonzalez and Holt, 2002, Roy et al., 2005, Schwager et al., 2006, Matthews and Gonzalez, 2007, Reuman et al., 2008, Roy and Holt, 2009, Schreiber, 2010]. To account for these autocorrelations, the environmental sequence of random variables, $\xi_1, \xi_2, \ldots$, can no longer be independent. However, they may be stationary or even asymptotically stationary i.e.

$$|P[X_t \in A_0, \ldots, X_{t+n} \in A_n] - P[X_{t+s} \in A_0, \ldots, X_{t+s+n} \in A_n]| \to 0 \text{ as } t \to \infty$$

for any Borel sets $A_0, \ldots, A_n \subset \Omega$, $n \geq 1$ and $s \geq 0$. Benaîm and Schreiber [2009] made this extension for structured populations models satisfying monotonicity assumptions i.e. Theorem 5.2. Similar extensions have yet to be made for multispecies models or non-monotonic structured models.

In order to apply the methods reviewed here, there is a desperate need for general methods to estimate the average per-capita growth rates $\lambda_i(\mu)$ and the dominant Lyapunov exponent $\gamma$. One approach is to consider “small noise approximations” of these quantities when distribution of $\xi_t$ is close to a Dirac measure. For structured population models where $A(0, \xi_t) = A + \varepsilon B_t$ for a fixed non-negative matrix $A$ and a sequence of random matrices $B_t$ with mean 0, Tuljapurkar [1990] developed second and higher order approximations of the dominant Lyapunov exponent $\gamma$ with respect to $\varepsilon$. This approximation yielded many useful insights into stochastic demography and metapopulation persistence [Wiener and Tuljapurkar, 1994, Boyce et al., 2006, Tuljapurkar and Haridas, 2006, Morris et al., 2008, Tuljapurkar et al., 2009, Schreiber, 2010].

Alternatively, for models of competing species, Chesson developed methods for estimating $\lambda_i(\mu)$ when the deterministic dynamics converge to an equilibrium and the small noise generates small demographic fluctuations around this equilibrium [Chesson, 1988, 1994, 2000]. Extending these methods to arbitrary species interactions and random perturbations from non-equilibrium dynamics, however, remains an important remaining challenge.

In conclusion, this review highlights the progress, challenges, and opportunities in using stochastic difference equations to understand the conditions necessary for population persistence. The speed at which this review becomes hopelessly outdated may be the best measurement of its success.

**References**


Department of Evolution and Ecology and the Center for Population Biology, University of California, Davis, California 95616

E-mail address: sschreiber@ucdavis.edu