



# A general theory of coexistence and extinction for stochastic ecological communities

Alexandru Hening<sup>1</sup> · Dang H. Nguyen<sup>2</sup> · Peter Chesson<sup>3</sup>

Received: 15 July 2020 / Revised: 17 February 2021 / Accepted: 12 April 2021 / Published online: 7 May 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

We analyze a general theory for coexistence and extinction of ecological communities that are influenced by stochastic temporal environmental fluctuations. The results apply to discrete time (stochastic difference equations), continuous time (stochastic differential equations), compact and non-compact state spaces and degenerate or non-degenerate noise. In addition, we can also include in the dynamics auxiliary variables that model environmental fluctuations, population structure, eco-environmental feedbacks or other internal or external factors. We are able to significantly generalize the recent discrete time results by Benaïm and Schreiber (J Math Biol 79:393–431, 2019) to non-compact state spaces, and we provide stronger persistence and extinction results. The continuous time results by Hening and Nguyen (Ann Appl Probab 28(3):1893–1942, 2018a) are strengthened to include degenerate noise and auxiliary variables. Using the general theory, we work out several examples. In discrete time, we classify the dynamics when there are one or two species, and look at the Ricker model, Log-normally distributed offspring models, lottery models, discrete Lotka–Volterra models as well as models of perennial and annual organisms. For the continuous time setting we explore models with a resource variable, stochastic replicator models, and three dimensional Lotka–Volterra models.

---

✉ Alexandru Hening  
alexandru.hening@tufts.edu

Dang H. Nguyen  
dangnh.maths@gmail.com

Peter Chesson  
pchesson@email.arizona.edu

<sup>1</sup> Department of Mathematics, Tufts University, Bromfield-Pearson Hall, 503 Boston Avenue, Medford, MA 02155, USA

<sup>2</sup> Department of Mathematics, University of Alabama, 345 Gordon Palmer Hall, Box 870350, Tuscaloosa, AL 35487-0350, USA

<sup>3</sup> Department of Ecology and Evolutionary Biology, The University of Arizona, Tucson, AZ, USA

**Keywords** Population dynamics · Environmental fluctuations · Stochastic difference equations · Stochastic differential equations · Environmental fluctuations · Auxiliary variables · Coexistence · Extinction

**Mathematics Subject Classification** 92D25 · 37H15 · 60H10 · 60J05 · 60J99 · 60F99

## Notation

$E_1$	The set of subsets of $\{1, \dots, n\}$ such that if $I \in E_1$ then the invariant measures living on $\mathcal{S}_+^I$ are attractors. See Definition 2.1
$E_2$	The set $E_2 := \mathcal{P}(n) \setminus E_1$
$F_i$	The fitness functions of species $i$
$G$	Function describing the dynamics of the auxiliary variable
$P_t$	The $t$ -transition operator or semigroup operator that acts on functions $H \in \mathcal{B}_b$ as $P_t(H)(\mathbf{z}) = \mathbb{E}_{\mathbf{z}}(H(\mathbf{Z}(t)))$
$P$	The transition operator of the discrete-time process $\mathbf{Z}(t)$
$V$	Lyapunov function coming up in Assumption 2.1 and in Assumption 3.1
$X_i(t)$	The density of species $i$ at time $t$
$\mathbf{E}, \mathbf{W}, \mathbf{B}$	Brownian motions on $\mathbb{R}^n, \mathbb{R}^{\kappa_0}, \mathbb{R}^{n+\kappa_0}$
$\mathbf{Z}(t)$	The process $(\mathbf{X}(t), \mathbf{Y}(t))$ at time $t$
$\text{Conv}(\mathcal{M})$	The set of invariant probability measures supported on $\mathcal{S}_0$
$\Delta$	The set $\Delta := \{\mathbf{x} \in \mathbb{R}_+^n \mid \sum_i x_i = 1\}$
$\mathbb{E}_{\mathbf{z}}$	The expectation under the probability measure $\mathbb{P}_{\mathbf{z}}$
$\mathbb{E}$	The expectation under the probability measure $\mathbb{P}$
$\Gamma, \Sigma, (\sigma_{ij})$	$\Gamma$ is a $(n + \kappa_0) \times (n + \kappa_0)$ matrix such that $\Gamma^\top \Gamma = \Sigma = (\sigma_{ij})_{(n+\kappa_0) \times (n+\kappa_0)}$ . The matrix $\Sigma$ encodes the covariance structure of the Brownian motions from the continuous time setting (3.1)
$\Gamma_{\mathbf{z}}$	The set $\Gamma_{\mathbf{z}} := \{\tilde{\mathbf{z}} \in \mathcal{S} \mid \tilde{\mathbf{z}} \text{ is accessible from } \mathbf{z}\}$ of points $\tilde{\mathbf{z}}$ such that for every neighborhood $U$ of $\tilde{\mathbf{z}}$ there is $t \geq 0$ for which $P_t(\mathbf{z}, U) > 0$
$\mathcal{M}^1$	The set of ergodic invariant probability measures supported on $\partial \mathbb{R}_+^n$ and which are attractors: $\mathcal{M}^1 := \{\mu \in \mathcal{M} : \mu \text{ satisfies Assumption 3.4}\}$ . This is in the continuous time setting without an auxiliary variable
$\mathcal{M}^2$	The set of ergodic invariant probability measures supported on $\partial \mathbb{R}_+^n$ and which are not attractors: $\mathcal{M}^2 := \mathcal{M} \setminus \mathcal{M}^1$ . This is in the continuous time setting without an auxiliary variable
$\mathcal{M}^I$	$\mathcal{M}^I := \{\mu \in \mathcal{M} \mid \mu(\mathcal{S}^I) = 1\}$ is the set of ergodic measures supported on the subspace $\mathcal{S}^I$
$\mathcal{M}^{I,+}$	$\mathcal{M}^{I,+} := \{\mu \in \mathcal{M} \mid \mu(\mathcal{S}_+^I) = 1\}$ be the set of ergodic measures supported on the subspace $\mathcal{S}_+^I$
$\mathcal{M}^{I,\partial}$	$\mathcal{M}^{I,\partial} := \{\mu \in \mathcal{M} \mid \mu(\mathcal{S}_0^I) = 1\}$ is the set of ergodic probability measures supported on $\mathcal{S}_0^I$
$\mathcal{M}$	The set of ergodic invariant probability measures supported on $\mathcal{S}_0$
$\mathbb{P}_{\mathbf{z}}$	$\mathbb{P}_{\mathbf{z}}(\cdot) = \mathbb{P}(\cdot \mid \mathbf{Z}(0) = \mathbf{z})$
$\mathbb{P}$	The probability measure
$\Pi_{t,\mathbf{z}}(B)$	The occupation measure $\Pi_{t,\mathbf{z}}(B) := \mathbb{E}_{\mathbf{z}} \tilde{\Pi}_t(B)$

$\mathbb{R}_+^\mu$	The set $\mathbb{R}_+^\mu := \{(x_1, \dots, x_n) \in \mathbb{R}_+^n : x_i = 0 \text{ if } i \in \mathcal{S}(\mu)^c\}$ . This is in the continuous time setting without an auxiliary variable
$\mathbb{R}_+^{\mu, \circ}$	The set $\mathbb{R}_+^{\mu, \circ} := \{(x_1, \dots, x_n) \in \mathbb{R}_+^n : x_i = 0 \text{ if } i \in \mathcal{S}(\mu)^c \text{ and } x_i > 0 \text{ if } x_i \in \mathcal{S}(\mu)\}$ . This is in the continuous time setting without an auxiliary variable
$\mathcal{S}(\mu)$	The species supported by the ergodic measure $\mu$ , i.e. $\mathcal{S}(\mu) = \{1 \leq i \leq n : \mu(\mathcal{S}^i) = 1\}$
$\mathcal{S}^I$	For a nonempty subset $I \subset \{1, \dots, n\}$ we define $\mathcal{S}^I := \{(\mathbf{x}, \mathbf{y}) \in \mathcal{S} \mid x_i = 0, i \notin I\}$ . $\mathcal{S}^I$ is the subspace in which all species not in $I$ are absent, and some or all species from $I$ are present. The set $\mathcal{S}^I$ represents a subcommunity where we can define persistence and extinction sets relative to that subcommunity
$\mathcal{S}^\emptyset$	Let $\mathcal{S}^\emptyset = \{(0, \mathbf{y}) \in \mathcal{S}\}$ be the set where all species are extinct
$\mathcal{S}^i$	The set $\mathcal{S}^i := \{\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in \mathcal{S} : x_i > 0\}$ is the subset of the state space where species $i$ persists
$\mathcal{S}_+^I$	If we restrict the process to $\mathcal{S}^I$ then the persistence set, where all species from $I$ persist, is given by $\mathcal{S}_+^I := \mathcal{S}^I \setminus \mathcal{S}_0^I$
$\mathcal{S}_+$	The persistence set $\mathcal{S}_+ := \mathcal{S} \setminus \mathcal{S}_0$
$\mathcal{S}_0^I$	If we restrict the process to $\mathcal{S}^I$ then the extinction set, where at least one species from $I$ is extinct, is given by $\mathcal{S}_0^I := \{\mathbf{z} \in \mathcal{S}^I \mid \prod_{j \in I} x_j = 0\}$
$\mathcal{S}_0$	The extinction set $\mathcal{S}_0 := \{(\mathbf{x}, \mathbf{y}) \in \mathcal{S} : \min_i x_i = 0\}$ , where at least one of the $n$ species is extinct
$\mathcal{S}_\eta$	$\mathcal{S}_\eta := \{(\mathbf{x}, \mathbf{y}) \in \mathcal{S} : \min_i x_i \leq \eta\}$
$\mathcal{S}$	The state space $\mathbb{R}_+^n \times \mathbb{R}^{k_0}$ of the process $\mathbf{Z}$
$\mathcal{U} = \mathcal{U}(\omega)$	The (random) set of weak*-limit points of $(\tilde{\Pi}_t)_{t \in \mathbb{N}}$
$\mathbf{c}^\top \mathbf{x}$	The scalar product $\mathbf{c}^\top \mathbf{x} = \sum_i c_i x_i$
$\delta_{\mathbf{z}}$	The Dirac mass at $\mathbf{z}$
$\gamma_1, \gamma_3, C, \rho$	Strictly positive constants coming up in Assumption 2.1
$\gamma_4, \gamma_5, C_4$	Strictly positive constants coming up in Assumption 3.1
$\mathbb{R}_+^n$	The positive orthant $[0, \infty)^n$
$\mathbb{Z}_+$	The set of positive integers
$\mathbf{X}(t)$	The species densities $(X_1(t), \dots, X_n(t))$ at time $t$
$\mathbf{Y}(t)$	The value of the auxiliary variable at time $t$
$\mathcal{B}_b$	The set $\mathcal{B}_b := \{H : \mathcal{S} \rightarrow \mathbb{R} \mid H \text{ Borel, bounded}\}$ of bounded Borel functions
$\mathcal{C}_b$	The set $\mathcal{C}_b := \{H : \mathcal{S} \rightarrow \mathbb{R} \mid H \text{ continuous, bounded}\}$ of bounded continuous functions
$\mathcal{L}$	The infinitesimal generator of the process $\mathbf{Z}(t)$ in the continuous time setting
$\mathcal{P}(n)$	The set of all subsets of $\{1, \dots, n\}$
$\partial \mathbb{R}_+^\mu$	The set $\partial \mathbb{R}_+^\mu := \mathbb{R}_+^\mu \setminus \mathbb{R}_+^{\mu, \circ}$ . This is in the continuous time setting without an auxiliary variable
$\phi$	A function that comes up in Assumption 2.2 which is needed for our extinction results

$\tilde{\Pi}_t(B)$	The random occupation measure of a Borel set $B$ : in discrete time $\tilde{\Pi}_t(B) := \frac{1}{t} \sum_{s=1}^t \delta_{\mathbf{Z}(s)}(B)$ and in continuous time $\tilde{\Pi}_t(B) = \frac{1}{t} \int_0^t \mathbf{1}_{\{\mathbf{Z}(s) \in B\}} ds$
$\xi(t)$	Random variable describing the environment at time $t$ (continuous time case) or on the interval $[t, t+1)$ (discrete time case)
$a \circ b$	$a \circ b := (a_1 b_1, a_2 b_2, \dots, a_n b_n)$
$f_i$	The drift of the dynamics of species $i$ is $X_i f_i(\mathbf{X}) dt$ in the continuous time setting
$g_i$	The diffusion term of the dynamics of species $i$ is $X_i g_i(\mathbf{X}) dE^i$ in the continuous time setting
$h$	The function $h(\mathbf{z}, \xi) = \left( \max_{i=1}^n \left\{ \max \left\{ F_i(\mathbf{z}, \xi), \frac{1}{F_i(\mathbf{z}, \xi)} \right\} \right\} \right)^{\gamma_3}$ coming up in Assumption 2.1
$r_i(\mu)$	The expected per-capita growth rate of species $i$ when introduced in the community described by $\mu$ . In the discrete-time case $r_i(\mu) = \int_{\mathcal{S}} \mathbb{E}[\log F_i(\mathbf{z}, \xi(1))] \mu(d\mathbf{z})$ , while in the continuous case $r_i(\mu) = \int_{\mathcal{S}} \left( f_i(\mathbf{z}) - \frac{\sigma_{ii} g_i^2(\mathbf{z})}{2} \right) \mu(d\mathbf{z})$

## Contents

1	Introduction	5
2	Discrete time results	6
2.1	Invariant measures and criteria for persistence	9
2.2	Extinction	14
2.3	Robust persistence and extinction	18
3	Continuous time results	19
3.1	Nondegenerate stochastic differential equations	23
4	Applications in discrete time	26
4.1	Dynamics in Markovian environments	26
4.1.1	One species	28
4.1.2	Two species	29
4.2	Structured populations	31
4.3	Environment-density interaction models in ecology	33
4.3.1	Perennial organisms: lottery models	34
4.3.2	Perennial organisms: Ricker recruitment variation	37
4.3.3	Perennial organisms: Ricker recruitment without covED	38
4.3.4	Annual plants with seed banks and predation	40
4.3.5	Additive models: Unbounded offspring distribution	41
4.3.6	Additive models: discrete Lotka–Volterra	41
5	Applications in continuous time	43
5.1	Structured populations	43
5.2	Continuous-time models with a resource variable	45
5.3	One species and one auxiliary variable	47
5.4	Stochastic replicator dynamics	47
5.5	Lotka Volterra: Two prey and one predator	51
6	Discussion	52
	Appendix A. Persistence proofs	55
	Appendix B. Extinction Proofs	64
	Appendix C. Robustness proofs	70
	Appendix D. Proof of Lemma 4.1	72

References . . . . .	74
----------------------	----

## 1 Introduction

Stochastic variation over time in the physical environment is an ever-present feature of natural ecosystems that has major effects on the lives of organisms (Chesson et al. 2013). This fact means that the dynamics of natural populations of plants and animals are not modeled well deterministically. Moreover, the usual deterministic attractors cannot reasonably characterize the asymptotic dynamics of natural communities. Although deterministic attractors, such as point equilibria, are often regarded as central tendencies about which populations fluctuate (May 1974), a stochastic environment can introduce qualitative changes in the nature of long-term dynamics that depart dramatically from the predictions of a deterministic attractor. For instance, in some cases, the deterministic attractor may predict species extinctions while the system in a stochastic environment predicts that all species persist (Chesson and Warner 1981; Li and Chesson 2016).

Prominent ecologists have long argued for stochastic models of the physical environment in ecological theory (Andrewartha and Birch 1954; Hutchinson 1961; Connell 1978; Strong 1986), but models of ecological dynamics in stochastic environments have posed serious challenges. Nevertheless, over the years, understanding has steadily improved. Of much significance is the recent development of general theory and techniques for establishing population persistence. Following early beginnings (Chesson 1982; Chesson and Ellner 1989; Ellner 1989), recent work (Schreiber 2011, 2012; Benaïm and Schreiber 2019; Hening and Nguyen 2018a) has provided key results for analyzing both discrete- and continuous-time variable environment models with multiple species. At first, these developments were restricted to the case of white noise environments (Schreiber 2011), which means that the vector of population sizes is a Markovian state variable. However, Benaïm and Schreiber (2019) goes beyond this case to allow auxiliary state variables permitting Markovian environmental variation, structured populations and feedbacks from population densities through other species. As remarkable as these developments are, they nevertheless have the serious limitation that the state space is required to be compact. Many ecological models naturally have non compact state spaces. Although it has been argued that finiteness of the universe, and finite capabilities of organisms, justify restriction to compact state spaces, modifying existing ecological models to compact state spaces requires unnatural assumptions, and restricts consideration to only a subset of ecological models that are useful in practice.

A key innovation of this work is removal of this compactness restriction. With this generalization, a number of ecological models that have led to new theoretical understanding can be fully rigorously analyzed. Moreover, boundedness restrictions on environmental variables in discrete-time models are removed, allowing a broader array of valuable ecological models to be analyzed. As a further generalization, in case not all species persist, we identify conditions for convergence to an invariant measure supported on the persistence set or the extinction set. Our results make use of the significant recent advances by Benaïm and Schreiber (2019), Benaïm (2018) and

Hening and Nguyen (2018a). In particular, we explicitly use some of the examples and computations by Schreiber (2011) and Benaïm and Schreiber (2019) in order to highlight how our analysis can strengthen earlier results as well as settle some previous conjectures. This work also highlights how the results apply to key ecological models that have been instrumental in showing strong effects of stochastic environmental variation on species coexistence. In this way, it strives to connect key ecological findings with the mathematics of stochastic persistence.

The paper is structured as follows. In Sect. 2 we present the framework and results in discrete time. The continuous time results appear in Sect. 3. Section 4 showcases how our results can be used to better understand specific discrete time ecological examples. Similarly, Sect. 5 applies the developed theory to continuous time examples. We discuss the relevance of our work in Sect. 6.

## 2 Discrete time results

We study a system of  $n$  species that interact nonlinearly and are influenced by temporal environmental variation. The species densities at time  $t \in \mathbb{Z}_+ := \{k \in \mathbb{Z} \mid k \geq 0\}$  are denoted by the vector  $\mathbf{X}(t) = (X_1(t), \dots, X_n(t))$  from  $\mathbb{R}_+^n := [0, \infty)^n$ . The species dynamics are influenced by and influence the auxiliary variable  $\mathbf{Y}(t)$ , which takes values in  $\mathbb{R}^{k_0}$ . This variable can correspond to eco-environmental feedbacks, forcing, the structure of each species or other factors. The dynamics of the species and of the feedback variable are changed by stochastic temporal variations. These environmental variations are described by the sequence of independent identically distributed random variables  $\xi(0), \xi(1), \dots$  which take values in a Polish space  $\Xi$ , that is, a separable completely metrizable topological space of which multidimensional Euclidean space is a key example. The state of the environment on the interval  $[t, t + 1)$  is described by  $\xi(t)$ , which then affects what the species densities will be at time  $t + 1$ . This is a deviation from a convention in some papers on this topic where the environment on  $(t, t + 1]$  is given as  $\xi(t + 1)$ , i.e. just a difference in the time index, which has no effect on the results. However, in our view this leads to more consistent terminology, and matches the ecological literature given that the species  $i$  has a fitness function  $F_i$  that at time  $t$  depends on the densities  $\mathbf{X}(t)$  of the various species, the auxiliary variable  $\mathbf{Y}(t)$ , and the environmental variable  $\xi(t)$ . Similarly, the auxiliary variable at time  $t + 1$  will be a function of  $\mathbf{X}(t)$ ,  $\mathbf{Y}(t)$  and  $\xi(t)$ . The process  $\mathbf{Z}(t) = (\mathbf{X}(t), \mathbf{Y}(t))$  tracks jointly the species densities and the auxiliary variable and lives on the state space  $\mathcal{S} := \mathbb{R}_+^n \times \mathbb{R}^{k_0}$ . As a result the dynamics will be

$$\begin{aligned} X_i(t + 1) &= X_i(t) F_i(\mathbf{X}(t), \mathbf{Y}(t), \xi(t)), \quad i = 1, \dots, n \\ \mathbf{Y}(t + 1) &= G(\mathbf{X}(t), \mathbf{Y}(t), \xi(t)). \end{aligned} \quad (2.1)$$

We define  $F(\cdot) = (F_1(\cdot), \dots, F_n(\cdot))$ .

**Assumption 2.1** The following assumptions are made throughout the paper:

- A1)** For each  $i = 1, \dots, n$ , the fitness function  $F_i(\mathbf{z}, \xi)$  is continuous in  $\mathbf{z} = (\mathbf{x}, \mathbf{y})$  for every  $\xi$ , where  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{y} = (y_1, \dots, y_{\kappa_0})$ , measurable in  $(\mathbf{z}, \xi)$ , and strictly positive.
- A2)** The function  $G : \mathbb{R}_+^n \times \mathbb{R}^{\kappa_0} \times \Xi \rightarrow \mathbb{R}^{\kappa_0}$  is continuous in  $\mathbf{z} = (\mathbf{x}, \mathbf{y})$  and measurable in  $(\mathbf{z}, \xi)$ .
- A3)** There exists a function  $V : \mathcal{S} \rightarrow \mathbb{R}_+$  and constants  $\gamma_1, \gamma_3, C > 0$  and  $\rho \in (0, 1)$  such that for all  $\mathbf{z} \in \mathcal{S}$
- i)  $|V(\mathbf{x}, \mathbf{y})| \geq |\mathbf{x}|^{\gamma_1} + 1$ ,
  - ii)  $\lim_{|\mathbf{z}| \rightarrow \infty} V(\mathbf{z}) = \infty$ , and
  - iii)  $\mathbb{E}[V(\mathbf{x} \circ F(\mathbf{z}, \xi(1)), G(\mathbf{z}, \xi(1)))h(\mathbf{z}, \xi(1))] \leq \rho V(\mathbf{z}) + C$  where  $a \circ b = (a_1 b_1, a_2 b_2, \dots, a_n b_n)$  is the Hadamard product,

$$h(\mathbf{z}, \xi) = \left( \max_{i=1}^n \left\{ \max \left\{ F_i(\mathbf{z}, \xi), \frac{1}{F_i(\mathbf{z}, \xi)} \right\} \right\} \right)^{\gamma_3}. \quad (2.2)$$

Assumptions A1) and A2) make sure that  $\mathbf{Z}(t)$  is a well-behaved Markov process. Measurability of  $F_i(\mathbf{z}, \xi)$  in  $(\mathbf{z}, \xi)$  is with respect to the natural Borel  $\sigma$ -algebra of the Polish space  $\Xi$ . Assumption A3) is also required in order to ensure that  $\mathbf{Z}(t)$  returns to compact subsets of  $\mathcal{S}$  exponentially fast and that the growth rates do not change too abruptly near infinity. A standard condition to ensure that  $\mathbf{Z}(t)$  returns to compact subsets of  $\mathcal{S}$  exponentially fast, which is critical to obtain exponential fast convergence to an invariant measure of a Markov chain, is A3)-ii) and a ‘weaker’ version of A3)-iii), namely:

$$\mathbb{E}[V(\mathbf{x} \circ F(\mathbf{z}, \xi(1)), G(\mathbf{z}, \xi(1)))] \leq \rho V(\mathbf{z}) + C.$$

This means that  $\mathbb{E}_{\mathbf{z}}(e^{r\tau}) < \infty$  for some  $r > 0$  where  $\tau$  is the first time the process enters a compact set. This finiteness is important in the study of Markov chains. Note that after visiting a compact set the process can leave this compact set for some time, before it returns again.

We next explain why condition A3)-iii) is needed. In order to better estimate the sample paths of the process, we want the growth rates to not change too abruptly. Since  $X_i(t+1) = X_i(t)F_i(\mathbf{X}(t), \mathbf{Y}(t), \xi(t))$ , to make sure that the growth rates do not change too much in one time step we want to be able to control  $X_i(t)F_i(\mathbf{X}(t), \mathbf{Y}(t), \xi(t))$ . We can do this if we know that

$$[F_i(\mathbf{X}(t), \mathbf{Y}(t), \xi(t))]^{1-\gamma_3} \leq F_i(\mathbf{X}(t), \mathbf{Y}(t), \xi(t)) \leq [F_i(\mathbf{X}(t), \mathbf{Y}(t), \xi(t))]^{1+\gamma_3}$$

for some  $\gamma_3 > 0$ . This is why  $h(\mathbf{z}, \xi)$  has the form (2.2) in A3)-iii). Finally, A3)-i) is needed because we need some boundedness of the moments of  $\mathbf{X}$ .

In general, if  $F_i$  is smaller than 1 when  $\mathbf{x}$  is large but bounded below in a certain sense, i.e.,  $\mathbb{E}\left(\frac{1}{F_i(\mathbf{x}, \xi)}\right)$  is bounded, then A3) will be satisfied.

**Remark 2.1** Note that in general it is not always easy to construct the Lyapunov function  $V$ . There is no definite answer for whether such a  $V$  exists. However, in many

applications, one can construct such a function. The examples presented throughout the paper show how one can do this. Furthermore, under stronger assumptions, for example, when the process enters and then stays forever in a compact set, the Lyapunov function is not required—see Remark 2.3.

**Remark 2.2** When we say that  $F_i(\mathbf{z}, \xi)$  is measurable in  $(\mathbf{z}, \xi)$  we note that the Polish space  $\Xi$  has a natural Borel  $\sigma$ -algebra.

**Remark 2.3** If it is known that the dynamics remain in a compact subset  $K \subset \mathbb{R}_+^n \times \mathbb{R}^{\kappa_0}$ , instead of A3), the following assumption is sufficient.

**A3C)** There exists a constant  $\gamma_3 > 0$  such that for all  $\mathbf{z} \in S$  we have

$$\mathbb{E}[h(\mathbf{z}, \xi(1))] < \infty$$

where

$$h(\mathbf{z}, \xi) = \left( \max_{i=1}^n \left\{ \max \left\{ F_i(\mathbf{z}, \xi), \frac{1}{F_i(\mathbf{z}, \xi)} \right\} \right\} \right)^{\gamma_3}.$$

**Remark 2.4** Assumption A3) (respectively A3C)) is strictly weaker than the assumptions made by Benaïm and Schreiber (2019). In particular Benaïm and Schreiber (2019) assume

**A3')** There is a compact subset  $K \subset \mathbb{R}_+^n \times \mathbb{R}^{\kappa_0}$  such that all solutions  $\mathbf{Z}(t)$  satisfy  $\mathbf{Z}(t) \in K$  for  $t \in \mathbb{Z}_+$  sufficiently large.

**A4')** For all  $i = 1, 2, \dots, n$ ,  $\sup_{\mathbf{z}, \xi} |\ln F_i(\mathbf{z}, \xi)| < \infty$ .

Note that assumption A3) need not imply that  $\mathbf{Z}(t)$  remain in a compact set, as examples from Sect. 4 attest. Note also that A3') and A4') imply A3C).

There are various ways in which one can define the persistence of the species from the studied ecosystem. We define the extinction set, where at least one species is extinct, by

$$S_0 := \{(\mathbf{x}, \mathbf{y}) \in S : \min_i x_i = 0\}$$

and the persistence set by

$$S_+ := S \setminus S_0.$$

For any  $\eta > 0$  let

$$S_\eta := \{(\mathbf{x}, \mathbf{y}) \in S : \min_i x_i \leq \eta\}$$

be the subset of  $S$  where at least one species is within  $\eta$  of extinction. We say (2.1) is *stochastically persistent in probability* (Chesson 1982) if for all  $\varepsilon > 0$  there exists  $\eta(\varepsilon) = \eta > 0$  such that for all  $\mathbf{z} \in S \setminus S_0$

$$\liminf_{t \rightarrow \infty} \mathbb{P}_{\mathbf{z}}\{\mathbf{Z}(t) \notin S_\eta\} > 1 - \varepsilon.$$



In essence, this means that if no species are extinct initially, with high probability their densities will not be found close to any of the axes at any given time in the future. We can also examine the realized frequencies with which the species have specific ranges of values for the period of time 1 to  $t$  with the “random occupation measure.” For any  $t \in \mathbb{N}$  define the *random occupation measure*

$$\tilde{\Pi}_t(B) := \frac{1}{t} \sum_{s=1}^t \delta_{\mathbf{Z}(s)}(B)$$

where  $\delta_{\mathbf{Z}(s)}$  is the Dirac measure at  $\mathbf{Z}(s)$  and  $B$  is any Borel subset of  $\mathcal{S}$ . We note that  $\Pi_t$  is a random probability measure and  $\Pi_t(B)$  tells us the proportion of time the system spends in  $B$  up to time  $t$ . We also define the occupation measure

$$\Pi_{t,z}(B) := \mathbb{E}_{\mathbf{z}} \tilde{\Pi}_t(B)$$

We say (2.1) is *almost surely stochastically persistent* (Schreiber 2012; Benaïm and Schreiber 2019) if for all  $\varepsilon > 0$  there exists  $\eta(\varepsilon) = \eta > 0$  such that for all  $\mathbf{z} \in \mathcal{S} \setminus \mathcal{S}_0$

$$\mathbb{P}_{\mathbf{z}} \left( \liminf_{t \rightarrow \infty} \tilde{\Pi}_t(\mathcal{S} \setminus \mathcal{S}_\eta) > 1 - \varepsilon \right) = 1.$$

## 2.1 Invariant measures and criteria for persistence

Following the seminal deterministic work by Hofbauer (1981) and the stochastic work by Chesson and Ellner (1989), Schreiber (2011), Benaïm (2018) and Hening and Nguyen (2018a) one needs to look at the Lyapunov exponents (expected per-capita growth rate) of certain invariant probability measures in order to give conditions for persistence. Intuitively, these invasion rates tell us if a species tends to increase or decrease when it is introduced into a subcommunity of species that is fluctuating according to a stationary probability distribution that it achieves asymptotically. We start by defining some of the required mathematical concepts.

The transition operator  $P : \mathcal{B}_b \rightarrow \mathcal{B}_b$  of the process  $\mathbf{Z}$  is an operator which acts on bounded Borel functions  $\mathcal{B}_b := \{H : \mathcal{S} \rightarrow \mathbb{R} \mid H \text{ Borel and bounded}\}$  as

$$PH(\mathbf{z}) = \mathbb{E}_{\mathbf{z}}[H(\mathbf{Z}(1))] := \mathbb{E}[H(\mathbf{Z}(1)) \mid \mathbf{Z}(0) = \mathbf{z}], \quad \mathbf{z} \in \mathcal{S}.$$

Thus, this operator gives conditional expectations of functions for changes of one time step. For multiple time steps, we can define for any  $t \in \mathbb{Z}_+$  the  $t$ -time transition operator  $P_t : \mathcal{B}_b \rightarrow \mathcal{B}_b$  via

$$P_t H(\mathbf{z}) = \mathbb{E}_{\mathbf{z}}[H(\mathbf{Z}(t))], \quad \mathbf{z} \in \mathcal{S}.$$

For any Borel set  $B \subset \mathcal{S}$  the function  $\mathbf{1}_B$  is the indicator function, which is 1 on  $B$  and 0 on the complement,  $B^c$ , of  $B$ . Sometimes we will write

$$P_t(\mathbf{z}, B) := P_t \mathbf{1}_B(\mathbf{z})$$

if  $B$  is a Borel set and  $\mathbf{z} \in \mathcal{S}$  any initial condition, thus defining  $t$ -step transition probabilities. We note that our assumptions imply that the process  $\mathbf{Z}(t)$  is Feller (Benaim 2018; Benaïm and Schreiber 2019): if  $C_b(\mathcal{S})$  is the set of continuous real valued functions defined on  $\mathcal{S}$ , then the mapping

$$(t, \mathbf{z}) \rightarrow P_t f(\mathbf{z})$$

is continuous. On the other hand, the operator  $P_t$  can also be used to obtain the probability distribution for  $\mathbf{Z}(t)$  starting with initial probability distribution (probability measure)  $\mu$  for  $\mathbf{Z}(0)$ . By duality, the operator  $P_t$  acts on Borel probability measures  $\mu$  by  $\mu \rightarrow \mu P_t$  where  $\mu P_t$  is the probability measure given by

$$\int_{\mathcal{S}} H(\mathbf{z})(\mu P_t)(d\mathbf{z}) := \int_{\mathcal{S}} P_t H(\mathbf{z})\mu(d\mathbf{z})$$

for all  $H \in C_b(\mathcal{S})$ , and is thus the probability distribution of  $\mathbf{Z}(t)$ .

A Borel probability measure  $\mu$  on  $\mathcal{S}$  is called an *invariant probability measure* if

$$\mu P_t = \mu, \quad t \geq 0.$$

The above equation tells us that if the system starts with a certain distribution  $\mu$  then it will have this distribution at any future time  $t \geq 0$ . Intuitively, this is the random analogue of a fixed point from dynamical systems.

The building blocks of the invariant probability measures are the *ergodic* invariant probability measures. These can be characterized as the extreme points of the set of all invariant probability measures. Equivalently, an invariant probability measure is ergodic if it cannot be written as a nontrivial convex combination of invariant probability measures. The set of ergodic invariant probability measures analogous to the set of stable points of a dynamical system. The invariant probabilities measures are then analogous to randomly starting the dynamical system at different stable points with given probabilities.

The invariant and ergodic probability measures are related to the long term behavior of the system. Suppose  $\mathbf{Z}(0)$  has distribution  $\mu$  and  $\mu$  is an invariant probability measure. One can then use Birkhoff's ergodic theorem to note that if  $H : \mathcal{S} \rightarrow \mathbb{R}$  is a measurable function that is integrable with respect to  $\mu$ , that is

$$\int_{\mathcal{S}} |H(\mathbf{z})| \mu(d\mathbf{z}) < \infty$$

then there exists a measurable integrable function  $\overline{H}$  such that, with probability one

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=0}^t H(\mathbf{Z}(s)) = \overline{H}(\mathbf{Z}(0))$$

The measure  $\mu$  is an ergodic probability measure if  $\overline{H}$  is a constant function for all bounded measurable functions  $h$  and one has, with probability one,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{s=0}^t H(\mathbf{Z}(s)) = \int_S H(\mathbf{z}) \mu(d\mathbf{z}).$$

Thus in the case of an ergodic initial distribution, the empirical observed averages are equal to the expectation, whereas with a nonergodic invariant initial measure, the empirical process follows one of the ergodic measures from which the initial distribution is constructed, but which of these is uncertain.

Note that these results only tell us about the long term behavior if  $\mathbf{Z}(0)$ 's initial distribution is an invariant probability measure. The methods developed below describe the long term behavior of the system for a wide class of ecological models, when  $\mathbf{Z}(0)$  is any positive initial condition.

For any species  $i$  let  $S^i := \{\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in S : x_i > 0\}$  be the subset of the state space for which this species has strictly positive density. If  $\mu$  is an ergodic measure and because  $S^i$  is an invariant set, i.e. if  $\mathbf{Z}(0) \in S^i$  then with probability 1 one has  $\mathbf{Z}(t) \in S^i, t \geq 0$ , we note that  $\mu(S^i) \in \{0, 1\}$ . This allows us to define the *species support* of  $\mu$  by  $S(\mu) = \{1 \leq i \leq n : \mu(S^i) = 1\}$ .

Suppose that we have a strict subcommunity of species at stationarity characterized by the ergodic measure  $\mu$ . Assume that at least one species will be extinct, so that  $S(\mu) \neq \{1, \dots, n\}$ . Let  $i \notin S(\mu)$  and introduce species  $i$  into the system at an infinitesimally low density. It turns out (Hofbauer 1981; Chesson and Ellner 1989; Chesson 1994; Benaïm and Schreiber 2019) that a key quantity is the *expected per-capita growth rate* of species  $i$  when introduced in the community described by  $\mu$

$$r_i(\mu) = \int_S \mathbb{E}[\log F_i(\mathbf{z}, \xi(1))] \mu(d\mathbf{z}). \quad (2.3)$$

This quantity is the invasion rate in the parlance of invasibility analysis and tells us whether a species  $i \notin S(\mu)$  tends to increase or decrease if introduced at a low density. In addition, if  $i \in S(\mu)$  then we can show that  $r_i(\mu) = 0$ , since the species supported by  $\mu$  already are at stationarity so they do not tend to grow or decrease exponentially fast. Denote by  $\mathcal{M}$  the set of all ergodic invariant probability measures supported on  $S_0$ . For a subset  $\tilde{\mathcal{M}} \subset \mathcal{M}$ , denote by  $\text{Conv}(\tilde{\mathcal{M}})$  the convex hull of  $\tilde{\mathcal{M}}$ , that is the set of probability measures  $\pi$  of the form  $\pi(\cdot) = \sum_{\mu \in \tilde{\mathcal{M}}} p_\mu \mu(\cdot)$  with  $p_\mu \geq 0, \sum_{\mu \in \tilde{\mathcal{M}}} p_\mu = 1$ . Using this notation  $\text{Conv}(\mathcal{M})$  is the set of all invariant probability measures supported on  $S_0$ .

**Remark 2.5** Note that one can show (see Benaïm 2018) that under some natural assumptions the set  $\text{Conv}(\mathcal{M})$  is convex and compact and  $\mu$  is ergodic if and only if it cannot be written as a nontrivial convex combinations of invariant probability measures. The ergodic decomposition theorem tells us that any invariant probability measure is a convex combination of ergodic measures. Furthermore, it can be shown that any two ergodic probability measures are either identical or mutually singular and that the topological supports of any mutually singular invariant measures are disjoint.

**Remark 2.6** In order to make the paper accessible to a wide variety of readers we will present our results without proofs in this section. The interested reader is asked to see Appendices A and B for the complete proofs of the various propositions and theorems.

The first result tells us that expected growth rates are well-defined and always zero for “resident species” in the parlance of invasibility analysis.

**Proposition 2.1** *Suppose  $\mu$  is an ergodic invariant measure. Then  $r_i(\mu)$  exists and is finite. Moreover,*

$$r_i(\mu) = 0, i \in \mathcal{S}(\mu).$$

Using this Proposition we can prove the following persistence result.

**Theorem 2.1** *Suppose that for all  $\mu \in \text{Conv}(\mathcal{M})$  we have*

$$\max_i r_i(\mu) > 0. \quad (2.4)$$

*Then (2.1) is almost surely stochastically persistent and stochastically persistent in probability.*

This result is a generalization to non compact state spaces of Theorem 1 by Benaïm and Schreiber (2019).

Although criterion (2.4) implies intuitively that species tend to increase when rare, and so are pushed away from the extinction set  $\mathcal{S}_0$ , some extra condition, which is supplied here by A3), is required, as the following example from Chesson (1982) shows. Stochastic persistence depends not only on the recovery of species from low densities but also on how violently a population can crash to low densities. Suppose  $N(t)$  is the population density of a small mammal population that has the dynamics of geometric growth up to a threshold  $K > 0$ , followed by random crashes from high density

$$N(t+1) = \begin{cases} 2N(t) & N(t) < K \\ \xi(t)N(t) & N(t) \geq K. \end{cases}$$

The environmental variable  $\xi(t)$  is assumed to take values in  $(0, 1)$  while  $\xi(1), \xi(2), \dots$  is an i.i.d. sequence. Let  $\delta_0$  be the Dirac mass at zero. This is the invariant measure that describes the ‘empty subcommunity’ where all species are extinct. Clearly

$$r(\delta_0) = \ln 2 > 0$$

and  $N(t)$  increases at low values no matter what the distribution of  $\xi(t)$  is. Moreover,  $\mathbb{P}(\lim_{t \rightarrow \infty} N(t) = 0) = 0$  because  $N(t)$  cannot only remain below  $K$  for finite periods of time. Nevertheless, there are situations in which  $N(t)$  converges to 0 in probability. For simplicity assume that  $\log_2 N(0), \log_2 \xi(t) \in \mathbb{Z}$ . This implies that  $\log_2 N(t)$  is a homogenous Markov chain on  $\mathbb{Z}$ . If  $\xi(t)$  is such that  $\mathbb{E} \log_2 \xi(t) = -\infty$ , then the process  $\log_2 N(t)$  is null recurrent and the expected return times to a given state are

infinite. Given that  $\log_2 N(t)$  is supported on  $(-\infty, 1 + \log_2 K]$ , it follows that for all  $x \in \mathbb{Z}$

$$\lim_{t \rightarrow \infty} \mathbb{P}(\log_2 N(t) > x) = 0$$

and therefore that  $N(t) \rightarrow 0$  in probability. As a result  $N(t)$  is not persistent in probability even though  $r(\delta_0) > 0$ .

The above example shows that one needs to ensure that the population will not violently crash to low densities. In the generality of our setting, this is accomplished by part A3) of Assumption 2.1.

**Remark 2.7** We note that the condition above is equivalent (Hening and Nguyen 2018a) to being able to find positive numbers  $p_1, \dots, p_n$  such that

$$\sum_{i=1}^n p_i r_i(\mu) > 0$$

for all  $\mu \in \mathcal{M}$ . This says that there exist weights of the species such that the weighted average of the expected per-capita growth rates is positive for all ergodic measures supporting a strict subset of the community. This criterion has first been introduced by Hofbauer (1981).

In order to prove stronger persistence results, we need some assumptions about the points of the state space the process can visit. We will follow the notation, methods and results developed by Meyn and Tweedie (1992) and Benaïm (2018). A point  $\tilde{\mathbf{z}} \in \mathcal{S}$  is said to be *accessible* from  $\mathbf{z} \in \mathcal{S}$  if for every neighborhood  $U$  of  $\tilde{\mathbf{z}}$  there exists  $t \geq 0$  such that  $P_t(\mathbf{z}, U) = P_t \mathbf{1}_U(\mathbf{z}) > 0$ . Define

$$\Gamma_{\mathbf{z}} := \{\tilde{\mathbf{z}} \in \mathcal{S} \mid \tilde{\mathbf{z}} \text{ is accessible from } \mathbf{z}\}$$

and for  $A \subset \mathcal{S}$

$$\Gamma_A = \bigcap_{\mathbf{z} \in A} \Gamma_{\mathbf{z}}.$$

Note that  $\Gamma_A$  is the set of points which are accessible from every point of  $A$ . We say a set  $A$  is *accessible* if for all  $\mathbf{z} \in \mathcal{S}_+$

$$\Gamma_{\mathbf{z}} \cap A \neq \emptyset.$$

**Theorem 2.2** *In addition to the assumptions of Theorem 2.1 suppose there exist  $\mathbf{z}^* \in \Gamma_{\mathcal{S}_+}$ , a neighborhood  $U$  of  $\mathbf{z}^*$ , a non-zero measure  $\iota$  on  $\mathcal{S}_+$ , and a probability measure  $\gamma$  on  $\mathbb{Z}_+$  such that for all  $\mathbf{z} \in U$*

$$\sum_{i=1}^{\infty} \gamma(i) P_i(\mathbf{z}, \cdot) \geq \iota(\cdot). \quad (2.5)$$

Then there exists a unique invariant probability measure  $\pi$  on  $\mathcal{S}_+$  and, with probability one, as  $t \rightarrow \infty$  the occupation measures  $(\Pi_{t,\mathbf{z}})_{t \in \mathbb{N}}$  converge weakly to  $\pi$  for any  $\mathbf{z} \in \mathcal{S}_+$ .

This result can be strengthened if the distribution of times,  $\gamma$ , can be replaced by a single time  $m^*$ , as follows.

**Theorem 2.3** Suppose that the assumptions of Theorem 2.2 hold, and there is a positive integer,  $m^*$ , for which

$$P_{m^*}(\mathbf{z}, \cdot) \geq \iota(\cdot)$$

for all  $\mathbf{z} \in U$ . Then there exists a unique invariant probability measure  $\pi$  on  $\mathcal{S}_+$  for which the distribution of  $\mathbf{Z}(t)$  converges in total variation to  $\pi$  as  $t \rightarrow \infty$  whenever  $\mathbf{Z}(0) = \mathbf{z} \in \mathcal{S}_+$ .

These results provide significant generalizations to Theorems 2 and 3 by Schreiber (2011) where only compact state spaces were considered.

**Remark 2.8** For a deeper understanding of the concepts of accessibility and irreducibility we refer the reader to the work by Meyn and Tweedie (1992) and Sections 4.2 and 4.3 by Benaim (2018). Having proven Theorem 2.1, the proofs of Theorems 2.2 and 2.3 follow from the well known convergence results of Meyn and Tweedie (1992) and Benaim (2018).

## 2.2 Extinction

A complete understanding of community dynamics requires not only understanding when species persist, but also when species go extinct. Mathematically, we are interested in conditions under which the process modelled by (2.1) converges to the boundary  $\mathcal{S}_0$ .

We need one extra condition, that ensures that the process can get close to the boundary, in order to get our extinction results.

**Assumption 2.2** There exists a function  $\phi : \mathcal{S} \rightarrow (0, \infty)$  and constants  $C, \delta_\phi > 0$  such that for all  $\mathbf{z} \in \mathcal{S}$

$$\mathbb{E}_{\mathbf{z}} V(\mathbf{Z}(1)) \leq V(\mathbf{z}) - \phi(\mathbf{z}) + C \quad (2.6)$$

and

$$\mathbb{E}_{\mathbf{z}} (V(\mathbf{Z}(1)) - PV(\mathbf{z}))^2 + \mathbb{E} |\log F(\mathbf{z}, \xi(1)) - \mathbb{E} \log F(\mathbf{z}, \xi(1))|^2 \leq \delta_\phi \phi(\mathbf{z}). \quad (2.7)$$

where  $V$  is the function satisfying Assumption 2.1.

**Remark 2.9** To obtain the extinction results, we need to manage the fluctuation of the process  $\mathbf{Z}(t)$ . This assumption bounds the variation of the process  $\mathbf{Z}(t)$ . We need that the next step position  $\mathbf{Z}(n+1)$  (resp.  $\log \mathbf{Z}(n+1)$ ) does not fluctuate too widely from the current position  $\mathbf{Z}(n)$  (resp.  $\log \mathbf{Z}(n)$ ). Mathematically,

we need certain boundedness of the quadratic terms  $\mathbb{E}_{\mathbf{z}} (V(\mathbf{Z}(1)) - PV(\mathbf{z}))^2$  and  $\mathbb{E} |\log F(\mathbf{z}, \xi(1)) - \mathbb{E} \log F(\mathbf{z}, \xi(1))|^2$ . Using well-known Lyapunov arguments, the boundedness can be obtained if the quadratic terms above are bounded by the dissipative part  $V(\mathbf{z}) - \mathbb{E}_{\mathbf{z}} V(\mathbf{Z}(1))$  (which is assumed to be bounded below, so it can be of the form  $\phi(\mathbf{z}) - C$  for some nonnegative function  $\phi(\mathbf{z})$ ).

In many cases, the function  $\phi(\mathbf{z})$  in (2.6) can be chosen to be  $V(\mathbf{z})$  or  $V^2(\mathbf{z})$ . In general the function  $\phi$  can be found by looking at  $V(\mathbf{z}) - \mathbb{E}_{\mathbf{z}} V(\mathbf{Z}(1))$ . Having Assumption 2.2, we can show the family of random occupation measures  $(\tilde{\Pi}_t)_{t \in \mathbb{N}}$  is tight. It is the discrete time analogue of Assumption 1.4 by Hening and Nguyen (2018a).

The first result tells us that under a certain condition there is asymptotic extinction with probability 1 no matter what the initial densities are.

**Theorem 2.4** *If there exist positive numbers  $p_1, \dots, p_n$  such that*

$$\sum_{i=1}^n p_i r_i(\mu) < 0$$

*for all ergodic probability measures supported by  $\mathcal{S}_0$  and  $\mathcal{S}_0$  is accessible then there exists  $\alpha > 0$  such that for any  $\mathbf{z} \in \mathcal{S}$*

$$\mathbb{P}_{\mathbf{z}} \left( \limsup_{t \rightarrow \infty} \frac{\ln d(\mathbf{Z}(t), \mathcal{S}_0)}{t} = -\alpha \right) = 1,$$

where  $d(\mathbf{z}, \mathcal{S}_0) = \min_i x_i$ .

The above theorem tells us when there is extinction but does not tell us exactly which species go extinct. In order to gain this extra information we need a few more definitions.

For a nonempty subset  $I \subset \{1, \dots, n\}$  define  $\mathcal{S}^I := \{(\mathbf{x}, \mathbf{y}) \in \mathcal{S} \mid x_i = 0, i \notin I\}$ . Note that  $\mathcal{S}^I$  is the subspace in which all species not in  $I$  are absent, and some or all species from  $I$  are present. The set  $\mathcal{S}^I$  represents a subcommunity where we can define persistence and extinction sets relative to that subcommunity.

Let  $\mathcal{S}^{\emptyset} = \{(0, \mathbf{y}) \in \mathcal{S}\}$  be the set where all species are extinct. If we restrict the process to  $\mathcal{S}^I$  then the extinction set, where at least one species from  $I$  is extinct, is given by  $\mathcal{S}_0^I := \{\mathbf{z} \in \mathcal{S}^I \mid \prod_{j \in I} x_j = 0\}$  and the persistence set, where all species from  $I$  persist, by  $\mathcal{S}_+^I := \mathcal{S}^I \setminus \mathcal{S}_0^I$ .

Let  $\mathcal{M}^I := \{\mu \in \mathcal{M} \mid \mu(\mathcal{S}^I) = 1\}$  be the set of ergodic measures supported on the subspace  $\mathcal{S}^I$ . A measure from  $\mathcal{M}^I$  will therefore support all, or a subset of, the species from  $I$ . Let  $\mathcal{M}^{I,+} := \{\mu \in \mathcal{M} \mid \mu(\mathcal{S}_+^I) = 1\}$  be the set of ergodic measures supported on the subspace  $\mathcal{S}_+^I$ . A measure from  $\mathcal{M}^{I,+}$  does not put any mass on the extinction set  $\mathcal{S}_0^I$ , i.e. all the species from  $I$  are present and no species is extinct. Finally, we define  $\mathcal{M}^{I,\partial} := \{\mu \in \mathcal{M} \mid \mu(\mathcal{S}_0^I) = 1\}$  to be the set of ergodic probability measures supported on  $\mathcal{S}_0^I$  respectively. If a measure is in  $\mathcal{M}^{I,\partial}$  this measure represents a subcommunity from  $I$  where at least one species is extinct.

Denote the (random) set of weak\*-limit points of  $(\tilde{\Pi}_t)_{t \in \mathbb{N}}$  by  $\mathcal{U} = \mathcal{U}(\omega)$ .

**Definition 2.1** Let  $E_1$  be the family of subsets  $I$  of  $\{1, \dots, n\}$  such that the following properties hold:

- (1) For all  $I \in E_1$  we have that  $\mathcal{M}^{I,+} \neq \emptyset$  and

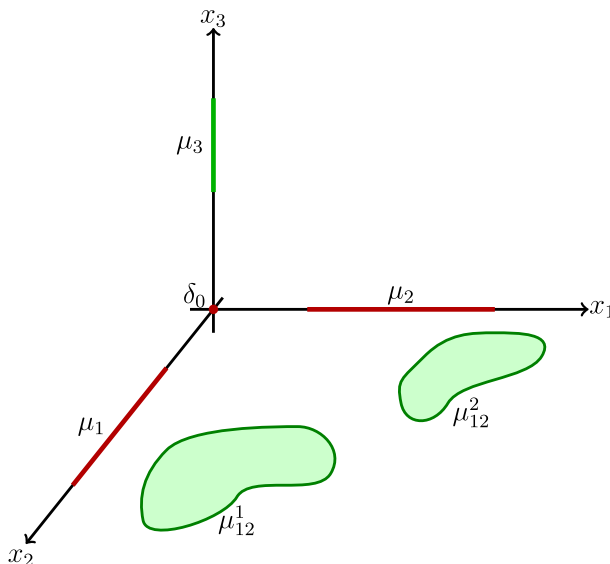
$$\max_{i \notin I} r_i(\mu) < 0, \mu \in \text{Conv}(\mathcal{M}^{I,+}). \quad (2.8)$$

- (2) If  $I \neq \emptyset$ , suppose further that for any  $\nu \in \text{Conv}(\mathcal{M}^{I,\partial})$ , we have

$$\max_{i \in I} r_i(\nu) > 0. \quad (2.9)$$

Let  $\mathcal{P}(n)$  be the set of all subsets of  $\{1, \dots, n\}$  and define  $E_2 := \mathcal{P}(n) \setminus E_1$ .

**Remark 2.10** The subset  $E_1$  tells us which subspaces  $S_+^I$  are attractors. If  $I \in E_1$  then invariant probability measures living on  $S_+^I$  are attractors while invariant measures living on the boundary  $S_0^I$  are repellers. This way, if  $\mathbf{Z}(t)$  gets close to  $S_+^I$  using (2.8) one can see the process gets attracted towards  $S_+^I$ . The second condition (2.9) ensures that the process gets repelled from the boundary  $S_0^I$  where the behaviour can be different. For some intuition regarding the various definitions see the example from Fig. 1.



**Fig. 1** Example of ergodic invariant measures. Green indicates attractors while red indicates repellers. In this example we picked:  $r_i(\delta_0) > 0, i = 1, 2, 3, r_2(\mu_1) > 0, r_1(\mu_2) > 0, r_1(\mu_3) < 0, r_2(\mu_3) < 0, r_3(\mu_{12}^1) < 0$  and  $r_3(\mu_{12}^2) < 0$ . This yields  $E_1 = \{\{1, 2\}, \{3\}\}$ ,  $\mathcal{S}(\mu_1) = \{1\}$ ,  $\mathcal{S}(\mu_2) = \{2\}$ ,  $\mathcal{S}(\mu_3) = \{3\}$ ,  $\mathcal{S}(\mu_{12}^1) = \{1, 2\}$ ,  $\mathcal{S}(\mu_{12}^2) = \{1, 2\}$ , the attractors living on the open  $x_1 - x_2$  positive quadrant are  $\mathcal{M}^{\{1,2\},+} = \{\mu_{12}^1, \mu_{12}^2\}$  and the attractor living on the open  $x_3$  positive axis is  $\mathcal{M}^{\{3\},+} = \{\mu_3\}$



**Theorem 2.5** For any  $I \in E_1$  not equal to  $\{1, \dots, n\}$ , there exists  $\alpha_I > 0$  such that, for any compact set  $\mathcal{K}^I \subset \mathcal{S}_+^I$ ,

$$\lim_{\text{dist}(\mathbf{z}, \mathcal{K}^I) \rightarrow 0, \mathbf{z} \in \mathcal{S}_+} \mathbb{P}_{\mathbf{z}} \left\{ \lim_{t \rightarrow \infty} \frac{\ln X_i(t)}{t} \leq -\alpha_I, i \notin I \right\} = 1.$$

**Biological Interpretation:** Suppose there exists at least one subspace  $\mathcal{S}_+^I$  for some  $I \subset \{1, \dots, n\}$  that is attracting. An attracting subspace is loosely speaking a subspace where every supported community at stationarity cannot be invaded by any outsider species. Then, if we start the system with densities  $\mathbf{z} \in \mathcal{S}_+$  which are close to  $\mathcal{S}_+^I$ , the species which are not supported by  $I$  go extinct exponentially fast with a probability that can be made arbitrarily close to 1.

**Theorem 2.6** Assume either that  $E_2 = \emptyset$  or that for all  $J \in E_2$

$$\max_i r_i(v) > 0, v \in \mathcal{M}^{J,+}.$$

If  $\bigcup_{I \in E_1} \mathcal{S}_+^I$  is accessible then

$$\sum_{I \in E_1} p_{\mathbf{z}, I} = 1$$

where

$$p_{\mathbf{z}, I} := \mathbb{P}_{\mathbf{z}} \left\{ \emptyset \neq \mathcal{U}(\omega) \subset \text{Conv}(\mathcal{M}^{I,+}) \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_j(t)}{t} \in \left\{ r_j(\mu) : \mu \in \text{Conv}(\mathcal{M}^{I,+}) \right\}, j \notin I \right\}.$$

If we assume that there exists a unique attracting ergodic measure in each attracting subspace we get a stronger result, which is an immediate corollary of the above.

**Theorem 2.7** Assume that every ergodic measure constructed as in Definition 2.1 is a unique attractor in its subspace, i.e. for any  $I \in E_1$  there exists a unique  $\mu_I$  satisfying (2.8) and (2.9). Furthermore, assume that  $E_2 = \emptyset$  or that for all  $J \in E_2$  and  $v \in \mathcal{M}^{J,+}$ ,  $\max_i r_i(v) > 0$ . Let  $\mathbf{z} \in \mathcal{S}_+^+$ . If  $\bigcup_{I \in E_1} \mathcal{S}_+^I$  is accessible then

$$\sum_{I \in E_1} p_{\mathbf{z}, I} = 1$$

where

$$p_{\mathbf{z}, I} := \mathbb{P}_{\mathbf{z}} \left\{ \mathcal{U}(\omega) = \{\mu_I\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_j(t)}{t} = r_j(\mu_I), j \notin I \right\}.$$

Furthermore, if for some  $I \in E_1$  the subspace  $\mathcal{S}_+^I$  is accessible one has  $p_{\mathbf{z}, I} > 0$ .

**Biological Interpretation:** Our result says the following. Suppose that the subspaces  $S_+^I$  for  $I \subset \{1, \dots, n\}$  are either of attracting or repelling type. An attracting subspace is loosely speaking a subspace where every supported community at stationarity cannot be invaded by any outsider species. A repelling subspace is a subspace where every supported community at stationarity can be invaded by at least one outsider species. If the attracting subspaces  $S_+^I$ ,  $I \in E_1$  are accessible, i.e. if starting at any population density one can get arbitrarily close to  $S_+^I$ , then with strictly positive probability the system converges to  $\mu_I$ . This means that the species given by  $I$  persist and converge to  $\mu_I$  while the species from  $I^c$  go extinct.

We note that the problem becomes much more complicated if there are subspaces that are neither attracting nor repelling. See Hening et al. (2021) for an example which involves rock-paper-scissors dynamics.

### 2.3 Robust persistence and extinction

Since all mathematical models are merely approximations of real ecological systems, it is important to study how the extinction and persistence results change under small perturbations of the analytical models. Ideally, small perturbations of the model should lead to similar long-term behavior. We next show that this is true in our setting as long as we add one natural assumption. Assume we perturb the dynamics of the system (2.1). We call  $\tilde{\mathbf{Z}}$  a  $\delta$ -perturbation of  $\mathbf{Z}$  if the dynamics is given by

$$\begin{aligned}\tilde{X}_i(t+1) &= \tilde{X}_i(t) \tilde{F}_i(\tilde{\mathbf{X}}(t), \tilde{\mathbf{Y}}(t), \xi(t)), \quad i = 1, \dots, n \\ \tilde{\mathbf{Y}}(t+1) &= \tilde{G}(\tilde{\mathbf{X}}(t), \tilde{\mathbf{Y}}(t), \xi(t))\end{aligned}\quad (2.10)$$

and

$$\sup_{\mathbf{z} \in \mathcal{S}} \sum_i \mathbb{E} \left[ \left\| \log \tilde{F}_i(\mathbf{x}, \mathbf{y}, \xi) - \log F_i(\mathbf{x}, \mathbf{y}, \xi) \right\| + \left\| \tilde{G}_i(\mathbf{x}, \mathbf{y}, \xi) - G_i(\mathbf{x}, \mathbf{y}, \xi) \right\| \right] \leq \delta. \quad (2.11)$$

Let  $\circ$  denote the element-wise product and  $\mathbf{1}_n$  be the vector in  $\mathbb{R}^n$  whose components are all 1. Assume that the function  $V$  from Assumption (2.1) satisfies the robust estimate

$$\mathbb{E} \left[ V(\mathbf{x}^\top (F(\mathbf{z}, \xi) \circ (\mathbf{1}_n + \tilde{\varepsilon}_1)), G(\mathbf{z}, \xi) + \tilde{\varepsilon}_2)) h(\mathbf{z}, \xi) \right] \leq \rho V(\mathbf{z}) + C, \quad \mathbf{z} \in \mathcal{S} \quad (2.12)$$

for all vectors  $\tilde{\varepsilon}_1, \tilde{\varepsilon}_2 \in \mathbb{R}^n$  such that  $|\tilde{\varepsilon}_1| \vee |\tilde{\varepsilon}_2| < \delta$ . We note that this estimate is trivially satisfied if the state space is compact.

**Theorem 2.8** *Assume the robust estimate (2.12) holds and the dynamics (2.1) satisfy the hypotheses of one of the persistence Theorems 2.1–2.3 or one of the extinction Theorems 2.4, 2.6 or 2.7. Then there exists  $\delta > 0$  such that every  $\delta$ -perturbation of (2.1) satisfies the same hypotheses and therefore has the same long term behavior (persistence or extinction).*

This generalizes the results from Proposition 2 by Schreiber (2011) to the setting when  $\mathbf{Z}$  is not restricted to a compact state space and the dynamics is influenced both by the species densities  $\mathbf{X}$  and by the auxiliary variables  $\mathbf{Y}$ .

### 3 Continuous time results

We continue to use the notation from Sect. 2, if not mentioned otherwise.

We work on a complete probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  with a filtration  $(\mathcal{F}_t)_{t \geq 0}$  satisfying the usual conditions. Consider the system

$$\begin{aligned} dX_i(t) &= X_i(t)f_i(\mathbf{X}(t), \mathbf{Y}(t))dt + X_i(t)g_i(\mathbf{X}(t), \mathbf{Y}(t))dE^i(t), \quad i = 1, \dots, n \\ dY_i(t) &= u_i(\mathbf{X}(t), \mathbf{Y}(t))dt + h_i(\mathbf{X}(t), \mathbf{Y}(t))dW_i(t), \quad i = 1, \dots, k \end{aligned} \quad (3.1)$$

taking values in  $\mathcal{S} = \mathbb{R}_+^n \times \mathbb{R}^{k_0}$  where  $\mathbb{R}_+^n := [0, \infty)^n$  and  $\mathbb{R}_+^{n, \circ} := (0, \infty)^n$ . We assume  $(\mathbf{E}(t), \mathbf{W}(t)) = (E^1(t), \dots, E^n(t), W_1(t), \dots, W_{k_0}(t))^T = \Gamma^\top \mathbf{B}(t)$  where  $\Gamma$  is a  $(n + k_0) \times (n + k_0)$  matrix such that  $\Gamma^\top \Gamma = \Sigma = (\sigma_{ij})_{(n+k_0) \times (n+k_0)}$  and  $\mathbf{B}(t) = (B_1(t), \dots, B_{n+k_0}(t))$  is a standard Brownian motion on  $\mathbb{R}^{n+k_0}$  which is adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . The SDE (3.1) is describing the dynamics of  $n$  interacting populations  $\mathbf{X}(t) = (X_1(t), \dots, X_n(t))_{t \geq 0}$ .

**Remark 3.1** We note that we can also treat compact state spaces for the species dynamics. That means we can study (3.1) when  $\mathbf{X}$  takes values in a compact state space, which for our ecological applications will usually be  $\Delta := \{\mathbf{x} \in \mathbb{R}_+^n \mid \sum_i x_i = 1\}$ . In this setting we assume the drift and the diffusion terms of  $\mathbf{X}$  are such that the dynamics remains on  $\Delta$ , or, in other words, that  $\Delta$  is an invariant set for the dynamics. See Schreiber (2011) for more examples of what the theory looks like in a compact state space.

Let  $\mathcal{L}$  be the infinitesimal generator of the process  $\mathbf{Z}$ . For smooth enough functions  $F : \mathbb{R}_+^n \times \mathbb{R}^{k_0} \rightarrow \mathbb{R}$  the generator  $\mathcal{L}$  acts as

$$\begin{aligned} \mathcal{L}F(\mathbf{z}) &= \sum_i x_i f_i(\mathbf{z}) \frac{\partial F}{\partial x_i}(\mathbf{z}) + \sum_i u_i(\mathbf{z}) \frac{\partial F}{\partial y_i}(\mathbf{z}) + \frac{1}{2} \sum_{i,j} \sigma_{ij} x_i x_j g_i(\mathbf{z}) g_j(\mathbf{z}) \frac{\partial^2 F}{\partial x_i \partial x_j}(\mathbf{z}) \\ &\quad + \frac{1}{2} \sum_{i,j} \sigma_{n+i, n+k_0} h_i(\mathbf{z}) h_j(\mathbf{z}) \frac{\partial^2 F}{\partial y_i \partial y_j}(\mathbf{z}) + \frac{1}{2} \sum_{i,j} \sigma_{i, n+j} x_i g_i(\mathbf{z}) h_j(\mathbf{z}) \frac{\partial^2 F}{\partial x_i \partial y_j}(\mathbf{z}). \end{aligned}$$

The following assumptions will be made in the continuous time setting.

**Assumption 3.1** The coefficients of (3.1) satisfy the following conditions:

- (1)  $f_i, g_i, u_i, h_i : \mathbb{R}_+^n \times \mathbb{R}^{k_0} \rightarrow \mathbb{R}$  are locally Lipschitz for any  $i = 1, \dots, n$ .
- (2) There exist numbers  $C_4, \gamma_4, \gamma_5 > 0$  and a function  $V \in C^2$  satisfying:

- $V(\mathbf{z}) \geq 1, \mathbf{z} \in \mathcal{S}$
- $\lim_{|\mathbf{z}| \rightarrow \infty} V(\mathbf{z}) = \infty$

- $V(\mathbf{z}) \geq |\mathbf{x}|^{\gamma_5}, \mathbf{z} = (\mathbf{x}, \mathbf{y}) \in \mathcal{S}$
- $$\mathcal{L}V(\mathbf{z}) \leq \left( C_4 - \gamma_4 \sum_{i=1}^n (1 + |f_i(\mathbf{z})| + |g_i(\mathbf{z})|^2) \right) V(\mathbf{z}), \mathbf{z} \in \mathcal{S}.$$

**Remark 3.2** Assumption 3.1 generalizes Assumption 1.1 in Hening and Nguyen (2018a). Here we assume the existence of a Lyapunov function to obtain a certain boundedness of the solution (in a moment sense). In many cases, the function  $V(\mathbf{z})$  will be of the form

$$V(\mathbf{z}) = 1 + \sum_{i=1}^n c_i x_i + \sum_{j=1}^{k_0} y_j^2.$$

When the diffusion is not too large and the drift points towards the origin as  $|\mathbf{z}|$  is large then the assumption is often satisfied. We note that in ecological terms this boils down to requiring that when species densities are large, there is a strong negative drift due to competition.

Under Assumption 3.1 one can show like in Lemma 3.1 from Hening and Nguyen (2018a) that  $\mathbf{Z}(t)$  is a Feller process that has pathwise unique solutions. Furthermore, if the process starts with  $\mathbf{Z}(0) \in \mathcal{S}_+$  then with probability 1 the process stays forever in  $\mathcal{S}_+$ .

One can associate to the Markov process  $\mathbf{Z}(t) = (\mathbf{X}(t), \mathbf{Y}(t))$  the semigroup  $(P_t)_{t \geq 0}$  defined by its action on bounded Borel measurable functions  $h : \mathcal{S} \rightarrow \mathbb{R}$  via

$$P_t h(\mathbf{z}) = \mathbb{E}_{\mathbf{z}}[h(\mathbf{Z}(t))], t \geq 0, \mathbf{z} \in \mathcal{S}.$$

The operator  $P_t$  can be seen to act by duality on Borel probability measures  $\mu$  by  $\mu \rightarrow \mu P_t$  where  $\mu P_t$  is the probability measure for which

$$\int_{\mathcal{S}} h(\mathbf{x})(\mu P_t)(d\mathbf{x}) := \int_{\mathcal{S}} P_t h(\mathbf{x})\mu(d\mathbf{x})$$

for all  $h \in C_b(\mathcal{S})$ . A probability measure  $\mu$  on  $\mathcal{S}$  is called invariant if  $\mu P_t = \mu$  for all  $t \geq 0$ . The invariant probability measure  $\mu$  is called ergodic if it cannot be written as a nontrivial convex combination of invariant probability measures. For any species  $i$  let  $\mathcal{S}^i := \{\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in \mathcal{S} : x_i > 0\}$  be the subset of the state space for which this species has strictly positive density. If  $\mu$  is an ergodic measure  $\mu(\mathcal{S}^i) \in \{0, 1\}$  we have defined the *species support* of  $\mu$  by  $\mathcal{S}(\mu) = \{1 \leq i \leq n : \mu(\mathcal{S}^i) = 1\}$ .

We can define in this setting for any  $t > 0$  the normalized random occupation measures

$$\tilde{\Pi}_t(\cdot) = \frac{1}{t} \int_0^t \mathbf{1}_{\{\mathbf{Z}(s) \in \cdot\}} ds,$$

the occupation measures

$$\Pi_{t,\mathbf{z}}(\cdot) = \mathbb{E}_{\mathbf{z}} \tilde{\Pi}_t(\cdot),$$

and the expected per-capita growth rates

$$r_i(\mu) = \int_S \left( f_i(\mathbf{z}) - \frac{\sigma_{ii} g_i^2(\mathbf{z})}{2} \right) \mu(d\mathbf{z}), i = 1, \dots, n.$$

Many of the discrete time results from Sect. 2 will hold in the continuous time setting. They represent generalizations of the work by Hening and Nguyen (2018a).

**Proposition 3.1** *Suppose  $\mu$  is an ergodic invariant measure. Then for any  $i \in I$  the quantity  $r_i(\mu)$  is well defined and finite. Moreover*

$$r_i(\mu) = \int_S \left( f_i(\mathbf{z}) - \frac{\sigma_{ii} g_i^2(\mathbf{z})}{2} \right) \mu(d\mathbf{z}) = 0, i \in \mathcal{S}(\mu).$$

**Remark 3.3** The above proposition provides a very powerful tool for computing  $r_\ell(\mu)$ ,  $\ell \notin \mathcal{S}(\mu)$  when  $f_i(\mathbf{z}) = \sum_j a_{ij} x_j$  and  $g_i(\mathbf{z}) = 1$ . In this setting we get the linear system

$$0 = r_i(\mu) = \sum_j a_{ij} \hat{x}_j - \frac{\sigma_{ii}}{2}, i \in \mathcal{S}(\mu)$$

where

$$\hat{x}_j = \int_S x_j \mu(d\mathbf{z}).$$

If this system has unique solutions we can then compute

$$r_\ell(\mu) = \sum_j a_{\ell j} \hat{x}_j - \frac{\sigma_{\ell\ell}}{2}, \ell \notin \mathcal{S}(\mu).$$

A point  $\tilde{\mathbf{z}} \in \mathcal{S}$  is said to be *accessible* from  $\mathbf{z} \in \mathcal{S}$  if for every neighborhood  $U$  of  $\tilde{\mathbf{z}}$  there exists  $t \geq 0$  such that  $P_t(\mathbf{z}, U) = P_t \mathbf{1}_U(\mathbf{z}) > 0$ . Define

$$\Gamma_{\mathbf{z}} := \{\tilde{\mathbf{z}} \in \mathcal{S} \mid \tilde{\mathbf{z}} \text{ is accessible from } \mathbf{z}\}$$

and for  $A \subset \mathcal{S}$

$$\Gamma_A = \bigcap_{\mathbf{z} \in A} \Gamma_{\mathbf{z}}.$$

Note that  $\Gamma_A$  is the set of points which are accessible from every point of  $A$ . We say a set  $A$  is *accessible* if for all  $\mathbf{z} \in \mathcal{S}_+$

$$\Gamma_{\mathbf{z}} \cap A \neq \emptyset.$$

We have the following sequence of results. These can be seen as the continuous time equivalents of the results from Sect. 2.

**Theorem 3.1** Suppose Assumption 3.1 holds and for all  $\mu \in \text{Conv}(\mathcal{M})$  we have

$$\max_i r_i(\mu) > 0.$$

Then (3.1) is almost surely stochastically persistent and stochastically persistent in probability.

Exactly like in Sect. 2 we define what it means for a set to be accessible or for the state space to be irreducible—see Meyn and Tweedie (1992) and Sections 4.2 and 4.3 by Benaïm (2018) for more details.

**Theorem 3.2** Suppose Assumption 3.1 holds and for all  $\mu \in \text{Conv}(\mathcal{M})$  we have

$$\max_i r_i(\mu) > 0.$$

In addition, assume there exist  $\mathbf{z}^* \in \Gamma_{S_+}$ , a neighborhood  $U$  of  $\mathbf{z}^*$ , a non-zero measure  $\iota$  on  $S_+$ , and a probability measure  $\gamma$  on  $\mathbb{R}_+$  such that for all  $\mathbf{z} \in U$

$$\int_0^\infty P_t(\mathbf{z}, \cdot) \gamma(dt) \geq \iota(\cdot). \quad (3.2)$$

Then, with probability one, the occupation measures  $(\Pi_{t,\mathbf{z}})$  converge weakly to a unique invariant probability measure  $\pi$  supported on  $S_+$  for any  $\mathbf{z} \in S_+$ .

**Theorem 3.3** Suppose that the assumptions of Theorem 3.2 hold, and there is a positive number,  $m^* > 0$ , for which

$$P_{m^*}(\mathbf{z}, \cdot) \geq \iota(\cdot)$$

for all  $\mathbf{z} \in U$ . Then there exists a unique invariant probability measure  $\pi$  on  $S_+$  and the distribution of  $\mathbf{Z}(t)$  converges in total variation to  $\pi$  exponentially fast whenever  $\mathbf{Z}(0) = \mathbf{z} \in S_+$ .

In order to prove extinction results we need to make the following extra assumption, which is the analogue of Assumption 2.2 from the discrete time setting.

**Assumption 3.2** There is  $C > 0$  such that

$$\left| \sum_{i=1}^n V_{x_i}(\mathbf{z}) x_i g_i(\mathbf{z}) + \sum_{i=1}^{K_0} V_{y_i}(\mathbf{z}) h_i(\mathbf{z}) \right| \leq C V(\mathbf{z}) \sum_{i=1}^n (1 + |f_i(\mathbf{z})| + |g_i(\mathbf{z})|^2), \mathbf{z} \in \mathcal{S}.$$

Suppose the notation is the same as the one from Sect. 2.

**Theorem 3.4** If  $E_1$  is nonempty, then for any  $I \in E_1$ , there exists  $\alpha_I > 0$  such that, for any compact extinction set  $\mathcal{K}^I \subset S_+^I$ , we have

$$\lim_{\text{dist}(\mathbf{z}, \mathcal{K}^I) \rightarrow 0, \mathbf{z} \in S^0} \mathbb{P}_{\mathbf{z}} \left\{ \lim_{t \rightarrow \infty} \frac{\ln X_i(t)}{t} \leq -\alpha_I, i \notin I \right\} = 1.$$

Let  $\mathcal{P}(n)$  be the set of all subsets of  $\{1, \dots, n\}$  and define  $E_2 := \mathcal{P}(n) \setminus E_1$ .

**Theorem 3.5** *Assume either that  $E_2 = \emptyset$  or that  $\max_i r_i(v) > 0$  for any  $v$  with  $v \in \mathcal{M}^{J,+}$  for some  $J \in E_2$ . If  $\bigcup_{I \in E_1} \mathcal{S}_+^I$  is accessible then*

$$\sum_{I \in E_1} p_{\mathbf{z}, I} = 1$$

where

$$p_{\mathbf{z}, I} := \mathbb{P}_{\mathbf{z}} \left\{ \mathcal{U}(\omega) \subset \text{Conv} \left( \mathcal{M}^{I,+} \right) \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_j(t)}{t} \in \left\{ r_j(\mu) : \mu \in \text{Conv} \left( \mathcal{M}^{I,+} \right) \right\}, j \notin I \right\}.$$

**Remark 3.4** We note that if  $(E_1, \dots, E_n)$  is non-degenerate, i.e. the matrix  $\Sigma$  is invertible, then each subspace  $\mathcal{S}_+^I$  is accessible.

If we assume that there exists a unique attracting ergodic measure in each attracting subspace we get a stronger result, which is an immediate corollary of the above.

**Theorem 3.6** *Assume that each  $I \in E_1$  is such that  $\mathcal{M}^{I,+} = \{\mu_I\}$ , i.e. for any  $I \in E_1$  there exists a unique  $\mu_I$  satisfying (2.8) and (2.9). Furthermore, assume that that  $E_2 = \emptyset$  or that  $\max_i r_i(v) > 0$  for any  $v \in \mathcal{M}^{J,+}$  for some  $J \in E_2$ . If  $\bigcup_{I \in E_1} \mathcal{S}_+^I$  is accessible then*

$$\sum_{I \in E_1} p_{\mathbf{z}, I} = 1$$

where

$$p_{\mathbf{z}, I} := \mathbb{P}_{\mathbf{z}} \left\{ \mathcal{U}(\omega) = \{\mu_I\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_j(t)}{t} = r_j(\mu_I), j \notin I \right\}.$$

Furthermore, if for some  $I \in E_1$  the subspace  $\mathcal{S}_+^I$  is accessible, then  $p_{\mathbf{z}, I} > 0$ .

The above theorems generalize the earlier persistence results by Schreiber (2011) and the persistence and extinction results by Hening and Nguyen (2018a).

### 3.1 Nondegenerate stochastic differential equations

We present the simplest setting, with no auxiliary variables, where all the accessibility and irreducibility conditions are satisfied. The dynamics of the  $n$  species is given by

$$dX_i(t) = X_i(t) f_i(\mathbf{X}(t)) dt + X_i(t) g_i(\mathbf{X}(t)) dE_i(t), \quad i = 1, \dots, n \quad (3.3)$$

taking values in  $[0, \infty)^n$ . Assume the following.

**Assumption 3.3** The coefficients of (3.1) satisfy the following conditions:

- (1)  $\text{diag}(g_1(\mathbf{x}), \dots, g_n(\mathbf{x}))\Gamma^\top \Gamma \text{diag}(g_1(\mathbf{x}), \dots, g_n(\mathbf{x})) = (g_i(\mathbf{x})g_j(\mathbf{x})\sigma_{ij})_{n \times n}$  is a positive definite matrix for any  $\mathbf{x} \in \mathbb{R}_+^n$ .
- (2)  $f_i(\cdot), g_i(\cdot) : \mathbb{R}_+^n \rightarrow \mathbb{R}$  are locally Lipschitz functions for any  $i = 1, \dots, n$ .
- (3) There exist  $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{R}_+^{n,\circ}$  and  $\gamma_b > 0$  such that

$$\limsup_{\|\mathbf{x}\| \rightarrow \infty} \left[ \frac{\sum_i c_i x_i f_i(\mathbf{x})}{1 + \mathbf{c}^\top \mathbf{x}} - \frac{1}{2} \frac{\sum_{i,j} \sigma_{ij} c_i c_j x_i x_j g_i(\mathbf{x}) g_j(\mathbf{x})}{(1 + \mathbf{c}^\top \mathbf{x})^2} + \gamma_b \left( 1 + \sum_i (|f_i(\mathbf{x})| + g_i^2(\mathbf{x})) \right) \right] < 0. \quad (3.4)$$

Note that part (1) of Assumption 3.3 means that the diffusion is nondegenerate. Meanwhile, (3.4) is the special case of Assumption 3.1, part (2) when there are no auxiliary variables and when the function  $V$  is chosen to be  $1 + \mathbf{c}^\top \mathbf{x}$ . As shown in Example 1.1 from Hening and Nguyen (2018a) condition (3.4) is satisfied for most ecological models (Schreiber 2011; Evans et al. 2015; Hening et al. 2018; Hening and Nguyen 2018a; Benaïm 2018). One can show (Hening and Nguyen 2018a) that Assumption 3.3 implies that all the irreducibility and accessibility conditions from Theorems 3.1–3.6 hold.

Consider  $\mu \in \mathcal{M}$ . Assume  $\mu \neq \delta^*$  where  $\delta^*$  is the Dirac mass at the origin, that is, at  $(0, \dots, 0)$ . Since the diffusion  $\mathbf{X}$  is non degenerate in each subspace, there exist  $0 < n_1 < \dots < n_k \leq n$  such that  $\text{supp}(\mu) = \mathbb{R}_+^\mu$  where

$$\mathbb{R}_+^\mu := \{(x_1, \dots, x_n) \in \mathbb{R}_+^n : x_i = 0 \text{ if } i \in \mathcal{S}(\mu)^c\}$$

for  $\mathcal{S}(\mu) := \{n_1, \dots, n_k\}$  and  $\mathcal{S}(\mu)^c := \{1, \dots, n\} \setminus \{n_1, \dots, n_k\}$ . If  $\mu = \delta^*$  then we note that  $\mathbb{R}_+^{\delta^*} = \{\mathbf{0}\}$ . Let

$$\mathbb{R}_+^{\mu,\circ} := \{(x_1, \dots, x_n) \in \mathbb{R}_+^n : x_i = 0 \text{ if } i \in \mathcal{S}(\mu)^c \text{ and } x_i > 0 \text{ if } x_i \in \mathcal{S}(\mu)\}$$

and  $\partial \mathbb{R}_+^\mu := \mathbb{R}_+^\mu \setminus \mathbb{R}_+^{\mu,\circ}$ . In this setting all the irreducibility and accessibility conditions hold and we have the following results.

**Theorem 3.7** Suppose that for all  $\mu \in \text{Conv}(\mathcal{M})$

$$\max_i \lambda_i(\mu) > 0.$$

Then there exists a unique invariant probability measure  $\pi$  on  $\mathbb{R}_+^{n,\circ}$  and the distribution of  $\mathbf{X}(t)$  converges in total variation to  $\pi$  exponentially fast whenever  $\mathbf{X}(0) = \mathbf{x} \in \mathbb{R}_+^{n,\circ}$ .

The following condition will imply extinction.

**Assumption 3.4** There exists a  $\mu \in \mathcal{M}$  such that

$$\max_{i \in \mathcal{S}(\mu)^c} \{\lambda_i(\mu)\} < 0. \quad (3.5)$$



If  $\mathbb{R}_+^\mu \neq \{\mathbf{0}\}$ , suppose further that for any  $\nu \in \text{Conv}(\mathcal{M}_\mu)$ , we have

$$\max_{i \in \mathcal{S}(\mu)} \{\lambda_i(\nu)\} > 0 \quad (3.6)$$

where  $\mathcal{M}_\mu := \{\nu' \in \mathcal{M} : \text{supp}(\nu') \subset \partial \mathbb{R}_+^\mu\}$ .

Define

$$\mathcal{M}^1 := \{\mu \in \mathcal{M} : \mu \text{ satisfies Assumption 3.4}\}, \quad (3.7)$$

and

$$\mathcal{M}^2 := \mathcal{M} \setminus \mathcal{M}^1. \quad (3.8)$$

To characterize the extinction of specific populations, we need some additional conditions.

**Assumption 3.5** Suppose that there is a  $\delta_1 > 0$  such that

$$\lim_{\|\mathbf{x}\| \rightarrow \infty} \frac{\|\mathbf{x}\|^{\delta_1} \sum_i g_i^2(\mathbf{x})}{1 + \sum_i (|f_i(\mathbf{x})| + |g_i(\mathbf{x})|^2)} = 0.$$

**Remark 3.5** This is a technical assumption which says that the noise terms do not grow too fast. We offer here some intuition. An application of Itô's Lemma yields that

$$\frac{\ln X_i(t)}{t} = \frac{\ln X_i(0)}{t} + \frac{1}{t} \int_0^t \left[ f_i(\mathbf{X}(s)) - \frac{g_i^2(\mathbf{X}(s))\sigma_{ii}}{2} \right] ds + \frac{1}{t} \int_0^t g_i(\mathbf{X}(s)) dE_i(s).$$

If  $\mathbf{X}$  is close to the support of an ergodic invariant measure  $\mu$  for a long time, then

$$\frac{1}{t} \int_0^t \left[ f_i(\mathbf{X}(s)) - \frac{g_i^2(\mathbf{X}(s))\sigma_{ii}}{2} \right] ds$$

can be approximated by the average with respect to  $\mu$

$$r_i(\mu) = \int_{\partial \mathbb{R}_+^n} \left( f_i(\mathbf{x}) - \frac{g_i^2(\mathbf{x})\sigma_{ii}}{2} \right) \mu(d\mathbf{x}).$$

while the term

$$\frac{\ln X_i(0)}{t}$$

is negligible as  $t \rightarrow \infty$ . We need Assumption (3.5) in order to guarantee that as  $t \rightarrow \infty$  we have with probability 1 that

$$\frac{1}{t} \int_0^t g_i(\mathbf{X}(s)) dE_i(s) \rightarrow 0.$$

To describe exactly which populations go extinct, we need an additional assumption which ensures that apart from those in  $\text{Conv}(\mathcal{M}^1)$ , invariant probability measures are “repellers”.

**Assumption 3.6** Suppose that one of the following is true

- $\mathcal{M}^2 = \emptyset$
- For any  $\nu \in \text{Conv}(\mathcal{M}^2)$ ,  $\max_i \lambda_i(\nu) > 0$ .

**Theorem 3.8** Suppose that Assumptions 3.3, 3.5 and 3.6 are satisfied and  $\mathcal{M}^1 \neq \emptyset$ . Then for any  $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$

$$\sum_{\mu \in \mathcal{M}^1} p_{\mathbf{x}}(\mu) = 1 \quad (3.9)$$

where

$$p_{\mathbf{x}}(\mu) := \mathbb{P}_{\mathbf{x}} \left\{ \mathcal{U}(\omega) = \{\mu\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_i(t)}{t} = \lambda_i(\mu) < 0, i \in \mathcal{S}(\mu)^c \right\} \\ > 0, \mathbf{x} \in \mathbb{R}_+^{n,\circ}, \mu \in \mathcal{M}^1.$$

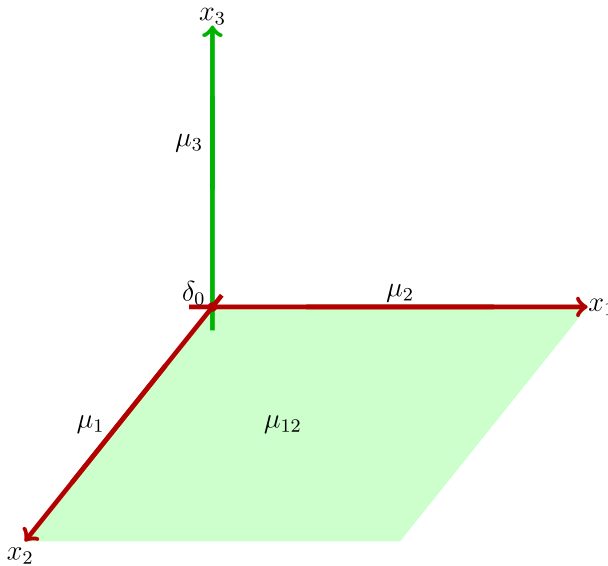
**Remark 3.6** The proofs of the continuous time SDE results from this section are very similar to the ones for the discrete time results from Sect. 2. They have therefore been omitted.

An example showcasing the various definitions is presented in Fig. 2.

## 4 Applications in discrete time

### 4.1 Dynamics in Markovian environments

As formulated in terms of equations (2.1), the discrete time system represents population dynamics subject to i.i.d environmental variation defined by the variable  $\xi(t)$ . However, the auxiliary variable  $\mathbf{Y}(t)$  might also be an environmental variable. Physical environmental variables in ecological models commonly reflect no feedback from the population variables, although in nature such feedback occurs but is not necessarily strong or easy to characterize. If  $\mathbf{Y}(t)$  has feedback only from itself, and not population densities, it becomes a Markov process and allows the Eq. (2.1) to represent



**Fig. 2** Example with attractors and repellers when the noise is nondegenerate. In this example we picked:  $r_i(\delta_0) > 0, i = 1, 2, 3, r_2(\mu_1) > 0, r_1(\mu_2) > 0, r_1(\mu_3) < 0, r_2(\mu_3) < 0$ , and  $r_3(\mu_{12}) < 0$ . This yields  $E_1 = \{\{1, 2\}, \{3\}\}$ ,  $\mathcal{S}(\mu_1) = \{1\}$ ,  $\mathcal{S}(\mu_2) = \{2\}$ ,  $\mathcal{S}(\mu_3) = \{3\}$ ,  $\mathcal{S}(\mu_{12}) = \{1, 2\}$ . Each subspace can have at most one ergodic measure. We have  $\mathcal{M}^{\{1,2\},+} = \{\mu_{12}\}$  and  $\mathcal{M}^{\{3\},+} = \{\mu_3\}$ . Note that the support of every ergodic measure is the whole subspace it lives in, for example:  $\mu_1$  lives on  $\mathbb{R}_+^{\mu_1, \circ} := \{(x_1, 0, 0) \in \mathbb{R}_+^3 : x_1 > 0\}$  and  $\mu_{12}$  lives on  $\mathbb{R}_+^{\mu_{12}, \circ} := \{(x_1, x_2, 0) \in \mathbb{R}_+^3 : x_1, x_2 > 0\}$

population dynamics in a Markovian environment as first suggested by Benaïm and Schreiber (2019). In this case the system (2.1) simplifies to

$$\begin{aligned} X_i(t+1) &= X_i(t)F_i(\mathbf{X}(t), \mathbf{Y}(t), \xi(t)), \\ \mathbf{Y}(t+1) &= G(\mathbf{Y}(t), \xi(t)). \end{aligned} \quad (4.1)$$

As formulated here, both  $\mathbf{Y}(t)$  and  $\xi(t)$  function as environmental variables, specifying the environment during the period  $[t, t+1)$ , which then influence the value of  $\mathbf{X}(t+1)$ . As a special case of (2.1), the general extinction and persistence theorems above continue to apply, although Assumptions 2.1 and 2.2 might be easier to check. We illustrate the issues that apply to Markovian environments with examples for single species, and pairs of interacting species. We begin with the assumption that  $\mathbf{Y}(t)$  has finitely many ergodic measures  $\mu_{\mathbf{Y}}^1, \dots, \mu_{\mathbf{Y}}^\ell$ .

For an individual species  $i$  we have

$$r_i(\delta_0 \times \mu_{\mathbf{Y}}^j) = \int_{\mathbb{R}^{\kappa_0}} \mathbb{E}[\ln F_i(0, \mathbf{y}, \xi(1))] \mu_{\mathbf{Y}}^j(d\mathbf{y}).$$

#### 4.1.1 One species

Species  $i$  will persist on its own if

$$r_i(\delta_0 \times \mu_{\mathbf{Y}}^j) > 0$$

for  $j = 1, \dots, \ell$ . This implies that at each ergodic probability measure of the environment, the log growth rate of species  $i$  at zero has to be positive. Once one checks which species will persist on their own, the next step is to look at two species systems and see which two species systems persist in isolation. We do this in the next section.

We note that if the Markovian environment  $\mathbf{Y}$  has only one ergodic measure  $\mu_{\mathbf{Y}}$  then computations are significantly simplified.

**Remark 4.1** Diaconis and Freedman (1999) give abstract conditions for when the environment

$$\mathbf{Y}(t+1) = G(\mathbf{Y}(t), \xi(t))$$

has a unique invariant measure. Suppose for each  $\xi$  the function  $G(\cdot, \xi)$  is Lipschitz with constant  $K_{\xi}$  and

- $\mathbb{E}[K_{\xi(1)}] < \infty$
- $\mathbb{E}|G(\mathbf{y}_0, \xi(1)) - \mathbf{y}_0| < \infty$  for some  $\mathbf{y}_0 \in \mathbb{R}^{k_0}$
- $\mathbb{E}[\ln K_{\xi(1)}] < \infty$

By Theorem 1.1 from Diaconis and Freedman (1999) the process  $\mathbf{Y}(t)$  has a unique invariant probability measure  $\mu_{\mathbf{Y}}$ . One example would be to assume that  $\mathbf{Y}(t+1) = A\mathbf{Y}(t) + \xi(t)$ . Then, if  $A$  has spectral radius less than one, the process  $\mathbf{Y}(t)$  converges (Diaconis and Freedman 1999) to a unique invariant probability measure  $\mu_{\mathbf{Y}}$ .

Suppose from now on that  $\mathbf{Y}$  has a unique ergodic measure  $\mu_{\mathbf{Y}}$ . In this case the dynamics is determined by

$$r_i(\delta_0 \times \mu_{\mathbf{Y}}) = \int_{\mathbb{R}^{k_0}} \mathbb{E} [\ln F_i(0, \mathbf{y}, \xi(1))] \mu_{\mathbf{Y}}(d\mathbf{y}).$$

If  $r_i(\delta_0 \times \mu_{\mathbf{Y}}) > 0$  we have persistence. If  $r_i(\delta_0 \times \mu_{\mathbf{Y}}) < 0$  the species goes extinct almost surely exponentially fast

$$\lim_{t \rightarrow \infty} \frac{\ln X(t)}{t} = r_X(\delta_0 \times \mu_{\mathbf{Y}}).$$

**Remark 4.2** One other way of modeling the environment would be to assume there are finitely many environmental states  $\{1, \dots, \ell\}$  and  $\mathbf{Y}(t)$  is an irreducible Markov chain with transition probabilities  $p_{ij} := \mathbb{P}[\mathbf{Y}(t+1) = j \mid \mathbf{Y}(t) = i]$ . The Markov chain will have a unique stationary distribution  $\nu := (\nu_1, \dots, \nu_{\ell})$ . In this setting the

only ergodic measure on  $\mathcal{S}_0$  is  $\delta_0 \times \nu$  and the persistence/extinction depends on the sign of

$$r_X(\delta_0 \times \nu) = \sum_{i=1}^{\ell} \nu_i \mathbb{E} [\ln F(0, i, \xi(1))].$$

If  $r_X(\delta_0 \times \nu) > 0$  we have persistence. If  $r_X(\delta_0 \times \nu) < 0$  the species goes extinct almost surely exponentially fast

$$\lim_{t \rightarrow \infty} \frac{\ln X(t)}{t} = r_X(\delta_0 \times \nu).$$

#### 4.1.2 Two species

Consider two interacting species that experience the effects of a Markovian environment

$$\begin{aligned} X_1(t+1) &= X_1(t) f_1(X_1(t), X_2(t), \mathbf{Y}(t), \xi(t)), \\ X_2(t+1) &= X_2(t) f_2(X_1(t), X_2(t), \mathbf{Y}(t), \xi(t)) \\ \mathbf{Y}(t+1) &= G(\mathbf{Y}(t), \xi(t)). \end{aligned} \quad (4.2)$$

We assume that  $\mathbf{Y}(t)$  has only one ergodic probability measure  $\mu_Y$ . Our results can be applied as follows. The first step is to check that Assumptions 2.1 and 2.2 hold—this is done once one knows more detailed properties of the system (noise and interaction terms). We exhibit how to check assumptions in specific examples in the next sections. Once the assumptions are checked, we look at the measure  $\delta_0 \times \mu_Y$

$$r_i(\delta_0 \times \mu_Y) = \int_{\mathbb{R}^{k_0}} \mathbb{E} [\ln f_i(0, \mathbf{y}, \xi(1))] \mu_Y(d\mathbf{y}). i = 1, 2.$$

If  $r_i(\delta_0 \times \mu_Y) > 0$  then species  $i$  survives on its own and the system converges to a unique invariant probability measure  $\mu_i$  supported on  $\mathcal{S}_+^i := \{\mathbf{x} \in \mathcal{S} \mid x_i \neq 0, x_j = 0, i \neq j\}$ . Remember that the (random) set of weak\*-limit points of the family of occupation measures  $(\tilde{\Pi}_t)_{t \in \mathbb{N}}$  is denoted by  $\mathcal{U} = \mathcal{U}(\omega)$ . Thus, if we say that  $\mathcal{U}(\omega) = \{\mu_1\}$ , this means that for the realization  $\omega$  we have  $\tilde{\Pi}_t \rightarrow \mu_1$  weakly.

- (i) Suppose  $r_1(\delta_0 \times \mu_Y), r_2(\delta_0 \times \mu_Y) > 0$ . The expected per-capita growth rates can be computed via

$$r_i(\mu_j) = \int_{(0, \infty) \times \mathbb{R}^{k_0}} \mathbb{E} [\ln f_i(x, \mathbf{y}, \xi(1))] \mu_j(dx d\mathbf{y}).$$

- If  $r_1(\mu_2) > 0$  and  $r_2(\mu_1) > 0$  we have coexistence and convergence of the distribution of  $\mathbf{X}(t)$  to the unique invariant probability measure  $\pi$  on  $\mathcal{S}_+$ .

- If  $r_1(\mu_2) > 0$  and  $r_2(\mu_1) < 0$  we have the persistence of  $X_1$  and extinction of  $X_2$ . In other words, for any  $\mathbf{x} \in \mathcal{S}_+$

$$\mathbb{P}_{\mathbf{x}} \left\{ \mathcal{U}(\omega) = \{\mu_1\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_2(t)}{t} = r_2(\mu_1) < 0, \right\} = 1.$$

- If  $r_1(\mu_2) < 0$  and  $r_2(\mu_1) > 0$  we have the persistence of  $X_2$  and extinction of  $X_1$ . In other words, for any  $\mathbf{x} \in \mathcal{S}_+$

$$\mathbb{P}_{\mathbf{x}} \left\{ \mathcal{U}(\omega) = \{\mu_2\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_1(t)}{t} = r_1(\mu_2) < 0, \right\} = 1.$$

- If  $r_1(\mu_2) < 0$  and  $r_2(\mu_1) < 0$  we have that for any  $\mathbf{x} \in \mathcal{S}_+$

$$p_{\mathbf{x},j} := \mathbb{P}_{\mathbf{x}} \left\{ \mathcal{U}(\omega) = \{\mu_j\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_i(t)}{t} = r_i(\mu_j) < 0, i \neq j \right\}$$

and

$$p_{\mathbf{x},1} + p_{\mathbf{x},2} = 1.$$

- (ii) Suppose  $r_1(\delta_0 \times \mu_{\mathbf{Y}}) > 0$ ,  $r_2(\delta_0 \times \mu_{\mathbf{Y}}) < 0$ . Then species 1 survives on its own and the system converges to the unique invariant probability measure  $\mu_1$  on  $\mathcal{S}_+^1$ .

- If  $r_2(\mu_1) > 0$  we have the persistence of both species and convergence of the distribution of  $\mathbf{X}(t)$  to the unique invariant probability measure  $\pi$  on  $\mathcal{S}_+$ .
- If  $r_2(\mu_1) < 0$  we have the persistence of  $X_1$  and the extinction of  $X_2$ . In other words, for any  $\mathbf{x} \in \mathcal{S}_+$

$$\mathbb{P}_{\mathbf{x}} \left\{ \mathcal{U}(\omega) = \{\mu_1\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_2(t)}{t} = r_2(\mu_1) < 0, \right\} = 1.$$

- (iii) Suppose  $r_1(\delta_0 \times \mu_{\mathbf{Y}}) < 0$ ,  $r_2(\delta_0 \times \mu_{\mathbf{Y}}) < 0$ . Then both species go extinct almost surely, i.e., for any  $\mathbf{x} \in \mathcal{S}_+$

$$\mathbb{P}_{\mathbf{x}} \left\{ \lim_{t \rightarrow \infty} \frac{\ln X_i(t)}{t} = r_i(\delta_0 \times \mu_{\mathbf{Y}}) < 0 \right\}, i = 1, 2.$$

These results are generalizations of those by Ellner (1989), Benaïm and Schreiber (2019) and Schreiber (2011). More specifically, we do not require any monotonicity or compactness, we show which species persist and which go extinct, and we also prove that there is always convergence to an ergodic probability measure supported by the persistence or extinction sets. Higher dimensional systems can be treated on a case by case basis. Usually, it is not possible to find the expected per-capita growth rates if the system is not of Lotka–Volterra type and the dimension is higher than 3. For a full classification in continuous time of three dimensional systems see Hening et al. (2021).

## 4.2 Structured populations

Populations of plants and animals generally have important internal structure. The most obvious structure is that the organisms consist of different ages. Importantly, different aged organisms can have different mortality rates, and different contributions to reproduction. Another way of structuring populations, which can be more helpful than age in some cases, is simply by life stage, for example as, egg, larva, and adult. In addition, populations do not live at points in space, but are spread over an area or region, with migration across the region. Invariably, the environments occupied by the population in different places differ somewhat in environmental characteristics, which can also be important to account for in the total growth of the population. When the region inhabited can be treated as discrete habitat patches, each with its own subpopulation of the species in question, the population structure is again discrete, just as stage, and age as an integer would be. There have been multiple studies of the interplay between dispersal and environmental heterogeneity (Hastings 1983; Chesson 1985, 2000a; Gonzalez and Holt 2002; Schmidt 2004; Roy et al. 2005; Schreiber 2010; Cantrell et al. 2012; Durrett and Remenik 2012). We are interested in models of  $m$  interacting populations that live in stochastic environments. The population is structured because each individual from population  $i$  can be in one of  $n_i$  individual states (these could be age, size or location). Then  $\bar{X}_i(t) = (X_{i1}(t), \dots, X_{in_i}(t))$  will be the row vector of population densities of individuals in different states for population  $i$  at time  $t \in \mathbb{N}$  and  $\bar{X}_i(t)$  will live on  $\mathbb{R}_+^{n_i}$ . The population state will be given by the row vector  $\bar{\mathbf{X}}(t) = (\bar{X}_1(t), \dots, \bar{X}_m(t))$  that will live on  $\mathbb{R}_+^n$  for  $n := \sum_{i=1}^m n_i$ . Just as before, the environment will be represented by the sequence of i.i.d random variables  $\xi(1), \xi(2), \dots, \xi(t), \dots$  where  $\xi(t)$  is the environment at the time step  $t$ . Assume that the environmental random variables are independent of the initial condition  $\bar{\mathbf{X}}(0)$  and take values in a probability space  $E$  that is equipped with a  $\sigma$ -field and a probability measure  $m$ . For all  $i$ , let  $A_i(\xi, \bar{\mathbf{X}}) = (a_i^{j,k}(\xi, \bar{\mathbf{X}}))$  be a non-negative  $n_i \times n_i$  matrix whose  $j - k$ th entry corresponds to the individuals moving from state  $j$  to state  $k$ . One can then write the population dynamics as

$$\bar{X}_i(t+1) = \bar{X}_i(t)A_i(\xi(t), \bar{\mathbf{X}}(t)). \quad (4.3)$$

These models have been studied in Roth and Schreiber (2014) and in a particular example in Benaïm and Schreiber (2019) under the restrictive assumption that the dynamics is constrained to a compact subset  $K \subset \mathbb{R}_+^n$ . Moreover, there are no general extinction results in Roth and Schreiber (2014) and Benaïm and Schreiber (2019).

We can transform (4.3) to our framework as follows. Define

$$X_i(t) := \sum_{j=1}^{n_i} X_{ij}(t)$$

to be the population size of species  $i$ ,  $\mathbf{X}(t) = (X_1(t), \dots, X_m(t))$ ,

$$Y_{ij}(t) = \frac{X_{ij}(t)}{X_i(t)}$$

the fraction of population  $i$  in state  $j$  and  $\mathbf{Y}_i(t) = (Y_{i1}(t), \dots, Y_{in_i}(t))$ ,  $\mathbf{Y}(t) = (\mathbf{Y}_1(t), \dots, \mathbf{Y}_m(t))$ . Then, simple calculations show that

$$\begin{aligned} X_i(t+1) &= X_i(t) \sum_{l,j} Y_{il} a_i^{j,l}(\xi(t), \mathbf{X}(t), \mathbf{Y}(t)), \quad i = 1, \dots, m \\ Y_{ij}(t+1) &= \frac{\sum_l Y_{il} a_i^{j,l}(\xi(t), \mathbf{X}(t), \mathbf{Y}(t))}{\sum_{l,j} Y_{il} a_i^{j,l}(\xi(t), \mathbf{X}(t), \mathbf{Y}(t))}, \quad i = 1, \dots, m \quad j = 1, \dots, n_i. \end{aligned} \quad (4.4)$$

Note that this system is in the correct form (4.3) and the state space is given by  $\mathcal{S} := \mathbb{R}_+^m \times \Delta_1 \times \dots \times \Delta_m$  where  $\Delta_i := \{\mathbf{x} \in \mathbb{R}_+^{n_i} : \sum_{j=1}^{n_i} x_j = 1\}$  is the simplex in  $\mathbb{R}^{n_i}$ .

The dynamics is well-defined on the extinction set  $\mathcal{S}_0 = \partial \mathbb{R}_+^m \times \Delta_1 \times \dots \times \Delta_m$ . For example, if  $X_1 = 0$  then the dynamics is

$$\begin{aligned} X_1(t+1) &= 0 \\ X_i(t+1) &= X_i(t) \sum_{l,j} Y_{il} a_i^{j,l}(\xi(t), 0, X_2(t), \dots, X_m(t), \mathbf{Y}(t)), \quad i = 2, \dots, m \\ Y_{ij}(t+1) &= \frac{\sum_l Y_{il} a_i^{j,l}(\xi(t), 0, X_2(t), \dots, X_m(t), \mathbf{Y}(t))}{\sum_{l,j} Y_{il} a_i^{j,l}(\xi(t), 0, X_2(t), \dots, X_m(t), \mathbf{Y}(t))}, \quad i = 1, \dots, m \quad j = 1, \dots, n_i. \end{aligned}$$

We will show how our theory can be used to generalize results by Roth and Schreiber (2014) and Benaïm and Schreiber (2019). Suppose there is only one species. We abuse notation and denote by  $(X_1(t), \dots, X_n(t))$  the densities of  $n$  patches and let  $X(t) = \sum_{i=1}^n X_i(t)$ ,  $Y_i(t) = \frac{X_i(t)}{X(t)}$  be the total population size and population fractions. We get the dynamics given by

$$\begin{aligned} X(t+1) &= X(t) \sum_{l,j} Y_l a_l^{j,l}(\xi(t), X(t), \mathbf{Y}(t)), \quad i = 1, \dots, m \\ Y_j(t+1) &= \frac{\sum_l Y_l a_l^{j,l}(\xi(t), X(t), \mathbf{Y}(t))}{\sum_{l,j} Y_l a_l^{j,l}(\xi(t), X(t), \mathbf{Y}(t))}, \quad j = 1, \dots, n \end{aligned} \quad (4.5)$$

with state space  $\mathcal{S} := \mathbb{R}_+ \times \Delta_n$  and extinction set  $\mathcal{S}_0 = \{0\} \times \Delta_n$ . One can show under certain assumptions that  $\mathbf{Y}(t)$  has a unique invariant measure  $\hat{\mu}$  on  $\mathcal{S}_0$ . We can use this to compute a growth rate  $r(\mu)$  such that the following theorem holds.

**Theorem 4.1** *Suppose that Assumption 2.1 holds. If  $r(\mu) > 0$  then  $\mathbf{X}(t)$  is stochastically persistent. In addition, under certain irreducibility conditions, there exist a unique invariant probability measure  $\pi$  and the law of  $\mathbf{X}(t)$  converges to  $\pi$  in total variation exponentially fast.*



If  $r(\mu) < 0$  and Assumption (2.2) holds, then for all  $\mathbf{x} \in \mathbb{R}_+$

$$\mathbb{P}_{\mathbf{x}} \left( \lim_{t \rightarrow \infty} X(t) = 0 \right) = 1.$$

### 4.3 Environment-density interaction models in ecology

A key question in community ecology, which the results here are ideally placed to address, is the role of temporal environmental fluctuations in species coexistence, i.e. the mutual persistence of interacting species. Models in community ecology have identified situations in which environmental fluctuations are essential for coexistence, or at least promote it, undermine coexistence, or have no effect on it (Chesson 1994). To systematize understanding, models of ecological community dynamics have been expressed in the following form,

$$N_j(t+1) = G_j(E_j(t), D_j(t))N_j(t), \quad (4.6)$$

with  $N_j(t)$  being the population density of species  $j$  at time  $t$ . The quantity  $G_j(E_j(t), D_j(t))$  is the multiplication rate, sometimes called the “finite rate of increase,” and corresponds to  $F_i(\mathbf{X}(t), \mathbf{Y}(t), \xi(t))$ , in the developments above. The two arguments,  $E_j(t)$  and  $D_j(t)$ , have special roles. The first of these,  $E_j(t)$ , is a component of the environmental variable  $\xi(t)$ , and is referred to as the environmental response of species  $j$ . The second variable,  $D_j(t)$ , is the species response to density-dependent processes. Thus,  $D_j(t)$  will reflect the densities of the interacting species even if it is not directly a function of them. For example,  $D_j(t)$  might be a function of resource shortage, or predator abundance, which could be auxiliary variables  $\mathbf{Y}(t)$  affected dynamically by the focal species densities, the  $N_j(t)$ . Originally,  $D_j(t)$  was conceived as reflecting competition, and designated  $C_j(t)$ , but the new notation,  $D_j(t)$ , is intended to indicate a broader class of species interactions including “apparent competition,” which is generated by density-dependent predation (Kuang and Chesson 2010; Chesson and Kuang 2010). By way of contrast, the environmental response does not reflect species densities, although it affects them. It is important to note, however, that  $D_j(t)$  will generally depend on the  $E_j(t)$ , but the slope of the relationship between  $D_j(t)$  and  $E_j(t)$  decreases to 0 as the density,  $N_j(t)$ , of species  $j$  approaches 0, a fact that has critical effects on species coexistence (Chesson 1994, 2019). The relationship between  $D_j(t)$  and  $E_j(t)$  is quantified by a statistical covariance between  $E_j(t)$  and  $D_j(t)$ , denoted  $\text{covED}_j$ . Positive  $\text{covED}_j$  means that favorable environmental conditions for species  $j$  are offset by unfavorable density effects limiting the gains in population that the species can have during favorable environmental conditions. However, when  $\text{covED}_j$  is evaluated using an invariant measure for which species  $j$  has zero density,  $D_j(t)$  will not be a function of  $E_j(t)$  and  $\text{covED}_j$  can be low or even negative depending on how much  $D_j(t)$  depends on the environmental responses of other species and how the environmental responses of the different species are correlated. These changes in  $\text{covED}_j$  for different invariant measures have critical effects on species coexistence provided one other condition applies, which is generally referred to as “buffered population growth” (Chesson 2019).

Buffered population growth is defined in terms of the quantity

$$\gamma_j = \frac{\partial^2 g_j(E, D)}{\partial E \partial D}, \quad (4.7)$$

where  $g_j(E_j(t), D_j(t)) = \ln G_j(E_j(t), D_j(t)) = \ln N_j(t+1) - \ln N_j(t)$ , defining population growth on the log scale. When  $\gamma_j$  is negative, population growth is said to be buffered because it means that high values of  $D_j(t)$  (unfavorable density effects) cause lower declines in  $g_j(E_j(t), D_j(t))$  when the environment is unfavorable too (low  $E_j(t)$ ). In effect, a double dose of unfavorable conditions is less than doubly unfavorable. As a consequence,  $r_j(\mu)$ , which is the same as the expected value of  $\ln G_j(E_j(t), D_j(t))$ , depends negatively on  $\text{covED}_j$ . The lower values of  $\text{covED}_j$  encountered for invariant measures,  $\mu$ , for which species  $j$  is extinct, thus favor positive  $r_j(\mu)$  values and hence persistence of species  $j$ . This is the “storage effect” species coexistence mechanism. It is quantified for the whole community (the “community average approach” (Chesson 2008; Yuan and Chesson 2015)) in terms of  $\gamma_j$  multiplied by differences between  $\text{covED}$  values for invariant measures  $\mu$  for which a species is extinct compared with those for which it is persistent. Importantly, the quantity  $\gamma_j$  is a reflection of the life-history properties of the organisms encoded in the model equations, and  $\text{covED}_j$  reflects how the direct ( $E_j(t)$ ) and indirect responses ( $D_j(t)$ ) of a species  $j$  to the environment are related to each other (Chesson and Huntly 1988). Below we study several different versions of this general model indicating how these key contributors to species coexistence emerge.

#### 4.3.1 Perennial organisms: lottery models

The lottery model of Chesson and Warner (1981) was the first model of competition between species in which stochastic temporal environmental variation emerged as a mechanism of species coexistence. It provides an example of the storage effect coexistence mechanism in operation (Chesson 1983, 1994). In this model, the individual organisms have exclusive sites or territories that provide the resources needed for the life of that individual. It is assumed that the space available for territories is strictly limiting. An individual organism must secure its own territory if it is to mature (“recruit to adulthood”) and reproduce, but once it has secured a site, it retains it for the rest of its life. We use a form of the model due to Chesson (1983). In this form, the variables  $N_i(t)$  are the adult densities of the species, and the environmental responses  $E_i(t)$  are per capita juvenile production rates, or in simple terms, “birth rates.” Thus,  $J_i(t) = E_i(t)N_i(t)$  is the number of juveniles of species  $i$  in year  $t$  seeking a site to mature as an adult. The relative ability of a juvenile to secure a site is given is a function,  $c_i(\mathbf{J}(t))$ , of the juvenile densities of all species, and so the fraction of available sites secured during year  $t$  by species  $i$  is

$$\frac{c_i(\mathbf{J}(t))J_i(t)}{\sum_{j=1}^n c_j(\mathbf{J}(t))J_j(t)}.$$

The organisms compete for space given up by adult death, which in year  $t$  is

$$\sum_{j=1}^n \delta_j(t) N_j(t).$$

where  $\delta_j(t)$  is the fraction of the adult population of species  $j$  that dies during  $(t, t + 1)$  and is assumed to be function of  $\xi(t)$ . As a consequence, the dynamics of the community are given by the following equations

$$N_i(t+1) = (1 - \delta_i(t))N_i(t) + \left( \sum_{j=1}^n \delta_j(t) N_j(t) \right) \left( \frac{c_i(\mathbf{J}(t))J_i(t)}{\sum_{j=1}^n c_j(\mathbf{J}(t))J_j(t)} \right), \quad i = 1, \dots, n. \quad (4.8)$$

To put this equation in the  $G(E, D)$  form (4.6), we can first define

$$D_i(t) = \frac{\sum_{j=1}^n c_j(\mathbf{J}(t))J_j(t)}{c_i(\mathbf{J}(t)) \sum_{j=1}^n \delta_j(t) N_j(t)}.$$

The numerator here is the total demand for space taking account of competitive ability, and the denominator is the competitive ability of species  $i$  times the supply of space, i.e.  $D_j(t)$  is “demand over supply” adjusted for relative competitive ability between species. The dynamical equation (4.8) now can be put in the form (4.6):

$$N_i(t+1) = \left( 1 - \delta_i(t) + \frac{E_i(t)}{D_i(t)} \right) N_i(t). \quad (4.9)$$

In this version of the lottery model, the function  $G_i$  will also vary with the environment if  $\delta_i(t)$  is temporally variable. Note that  $\sum_{j=1}^n N_j(t)$  remains constant for all  $t \in \mathbb{Z}_+$  and is equal to the total area available for individual sites. We set this total space equal to 1 by choice of units. Therefore the dynamics take place on the compact state space  $\Delta := \{\mathbf{x} \in \mathbb{R}_+^n \mid \sum_i x_i = 1\}$ . Note that on a compact state space Assumptions 2.1 and 2.2 are automatically satisfied.

If there are only two species, variations on this model have been analyzed in detail by Chesson (1982), Chesson (1983) and Schreiber (2011). Note that the extinction set is  $\Delta_0 = \{(1, 0), (0, 1)\}$ . As such, we only have to compute the per capita growth rates of the Dirac masses  $\delta_{(1,0)}, \delta_{(0,1)}$ . This is trivial and yields

$$\lambda_1 := r_1(\delta_{(0,1)}) = \mathbb{E} \left[ \ln \left( 1 - \delta_1(1) + \delta_1(1) \frac{c_1(0, E_2(0))E_1(0)}{c_2(0, E_2(0))E_2(0)} \right) \right]$$

and

$$\lambda_2 := r_2(\delta_{(1,0)}) = \mathbb{E} \left[ \ln \left( 1 - \delta_2(1) + \delta_2(1) \frac{c_2(E_1(0), 0)E_2(0)}{c_1(E_1(0), 0)E_1(0)} \right) \right]$$

It is not hard to see that the irreducibility and accessibility assumptions are satisfied for many choices of  $c_i$ ,  $\delta_i$  and  $E_i$ . For example, it is sufficient for  $\delta_i(0) > 0$ , a.s., and the ratio  $c_1(\mathbf{J}(0))E_1(0)/c_2(\mathbf{J}(0))E_2(t)$  to have a positive continuous probability density function over the domain  $(0, \infty)$ , conditional on the  $\delta_i(0)$  and  $\mathbf{N}(0)$ . Then next result presents generalizations of Chesson (1982, Theorems 3.5, 5.1 and 5.2).

**Theorem 4.2** *The following hold:*

- If  $\lambda_i > 0$ ,  $i = 1, 2$  then the system is stochastically persistent and furthermore the distribution of  $(N_1(t), N_2(t))$  converges to a unique invariant probability measure  $\pi$  on  $\Delta \setminus \Delta_0 = \{(x_1, x_2) \in \mathbb{R}_+^2, x_1, x_2 > 0, x_1 + x_2 = 1\}$ .
- If  $\lambda_i > 0$  and  $\lambda_j < 0$  then  $\mathbb{P}(N_i(t) \rightarrow 1, N_j(t) \rightarrow 0) = 1$ .
- If  $\lambda_1 < 0$  and  $\lambda_2 < 0$  then  $\mathbb{P}(N_1(t) \rightarrow 1 \text{ or } N_2(t) \rightarrow 0) = 1$  and  $\mathbb{P}(N_i(t) \rightarrow 1) > 0$ ,  $i = 1, 2$ .

Approximations to these low-density growth rates,  $\lambda_i$ , have been developed in number of publications for both the two-species and multispecies cases (Chesson 1989, 1994, 2003). Most important, the approximations reveal the biological circumstances leading to the three different possibilities in Theorem 4.2. In the case where the functions  $c_j$  are simply constants, there is no possibility of coexistence unless the environment varies. Most important, sufficient variation in the ratio  $E_1/E_2$  about the value 1, while the  $\mathbb{E}[\ln(1 - \delta_i(0))]$  are bounded from  $-\infty$ , guarantees coexistence, i.e. stochastic persistence of both species (Chesson and Warner 1981). This example is important as the first demonstration of how a stochastically varying environment could promote species coexistence. This ability was subsequently traced to a negative value for  $\gamma_i$  and covED (Chesson and Huntly 1988; Chesson 1994). This negative value of  $\gamma_i$  means that the growth of the population is less sensitive to the density-dependent variable  $D_i(t)$ , in this case competition, when environmental conditions for recruitment, as given by  $E_i(t)$ , are poor. The outcome is that the population can increase strongly during environmentally favorable times without suffering catastrophic losses from competition under environmentally poor conditions. This property defines the storage component of the storage effect coexistence mechanism. This feature arises quite generally from common life-history properties of natural populations, in particular, consistently high adult survival while recruitment to the adult population is highly variable (Chesson and Huntly 1988).

The full storage effect mechanism combines this property with covED, which explains why a species at a lower density gains boosts to population growth by having a higher frequency of times when  $E_i$  is high while  $D_i$  is low. For example, for constant  $\delta$ 's, and species 1 present alone ( $1 \in S(\mu)$ ,  $2 \notin S(\mu)$  in a two-species setting), covED<sub>1</sub> is positive because then  $D_1 = E_1/\delta_1$ . However, in this setting  $D_2 = (c_1(E_1, 0)/c_2(E_1, 0))E_1\delta_1$ , which need not be strongly correlated with  $E_2$  leading to a low covED<sub>2</sub> for species 2. As a result of this low covED, species 2 can have times when it is strongly favored by the environment while experiencing low competition. The fact that it also has times when it is disfavored by the environment and experiences high competition is not so important due to the negative value of  $\gamma_1$ . Species 1, however, has no such advantages due to its high covED, and so there is an overall net gain to species 2 in this setting, and generally to a species at low

density invading the population of another species. This is the storage effect mechanism, and explains how the variable environment boosts the  $\lambda_i$ 's and thereby promotes coexistence in the lottery and other models, as discussed further below.

These coexistence results depended on highly variable juvenile production rates as the choice for  $E_i(t)$ . In contrast, if adult survival is made highly variable and  $1 - \delta_i(t)$  is chosen as  $E_i(t)$ ,  $\gamma_i$  is positive, making the population growth rate of a species more sensitive to competition under unfavorable environmental conditions. As a consequence, increasing variation in  $(1 - \delta_1(t))/(1 - \delta_2(t))$  about 1 leads to the third possibility in Theorem 4.2, where  $\lambda_1$  and  $\lambda_2$  are both negative, and one of the species must go extinct, without the identity of that species being predictable ("random exclusion") due to a negative storage effect from positive  $\gamma_i$  (Chesson and Warner 1981). Finally, the second case, where  $\lambda_i > 0$  and  $\lambda_j < 0$ , giving certain persistence of one species, and certain extinction of the other, occurs when  $\delta_1 = \delta_2 = 1$ , regardless of the magnitude of environmental variation. In this case  $\gamma_i = 0$ , and there is no storage effect, either positive or negative. Each of the three outcomes in Theorem 4.2 is also possible without environmental variation but due to dependence of the  $c_i(\mathbf{J}(t))$  on the juvenile densities  $\mathbf{J}(t)$ . For example, if this dependence specifies that intraspecific effects are stronger than interspecific effects, then coexistence occurs, as is well-understood in deterministic models (Chesson 2018). However, the key interest here is how the stochastic environment can create these outcomes.

#### 4.3.2 Perennial organisms: Ricker recruitment variation

Like the lottery model, this model has long-lived adults, and juveniles that compete for the opportunity to recruit into the adult population, but in this case competition is not for space, but instead for a specific resource of richness  $S(t)$  that fluctuates over time with the environment. Competition takes the Ricker form (Chesson 1994), and the equations can be written

$$N_i(t+1) = N_i(t) \left(1 - \delta_i + S(t)e^{E_i(t)-D(t)}\right) \quad (4.10)$$

with

$$D(t) = \sum_{j=1}^n \alpha_j e^{E_j(t)} N_j(t). \quad (4.11)$$

Here  $E_j(t)$  can be interpreted as the logarithm of the per capita number of births in the interval  $(t, t+1)$ . The density response variable,  $D(t)$ , is the same for all species in this case, and so is not subscripted. This model also takes the  $G(E, D)$  form (4.6) if  $S(t)$  is a constant. More generally, this form can be retained if  $E_j(t)$  is replaced by  $E_j(t) + \ln S(t)$ , but in the definition of  $D$  this extra component,  $\ln S(t)$ , would need to be subtracted from  $E_j(t)$  to preserve the model. Most important, like the lottery model with variable juvenile production, in this model the interaction coefficient  $\gamma_i$  is negative, and covED is present due to the dependence of  $D_j(t)$  on the  $E_j(t)$ . These features allow environmental variation to enable coexistence by the storage effect

(Chesson 1994, 2003). In the absence of environmental variation, coexistence at a stable equilibrium is not possible due to the fact that all species share the same density response variable  $D(t)$  (Chesson and Huntly 1997). However, sustained deterministic fluctuations are a possibility in this model, and can support coexistence under some circumstances by a mechanism termed “relative nonlinearity” (Kuang and Chesson 2008). The theory presented here, however, enables rigorous demonstration of when coexistence can occur in a stochastic environment.

The next lemma gives conditions under which the general assumptions hold for this class of models.

**Lemma 4.1** *Suppose  $\{(E_i(t), S(t))_{i=1,\dots,n}\}$ ,  $t \in \mathbb{N}$  is a sequence of  $n+1$ -dimensional random variables, i.i.d. over  $t$  such that  $\mathbb{E}[S(t)e^{E_j(t)}]^2 < \infty$ . Then the model given by (2.4) and (4.11) satisfies Assumption 2.1 by taking a small enough  $\gamma_3 > 0$ , and*

$$V(\mathbf{z}) = \sum_j z_j + 1.$$

Assumption 2.2 holds with

$$\phi(\mathbf{z}) = \delta V(\mathbf{z})$$

for some  $\delta > 0$ . Moreover, if the support of  $\ln S(t) + \max_j (E_j(t) \ln \delta_j)$  contains values less than 0 then the boundary is accessible.

**Remark 4.3** We have chosen to work with  $E_j(t)$  as well as  $S(t)$  as part of the environment  $\xi$ . So in this case there is no  $\mathbf{Y}$  auxiliary variable,  $\mathbf{z} = \mathbf{x}$ ,  $\mathbf{X} = (N_1(t), \dots, N_n(t))$  and  $\xi(t) := (E_1(t), \dots, E_n(t), S(t))$ . We could instead treat  $S(t)$  as an auxiliary variable—the results would be unchanged.

**Proof** See Appendix 1. □

#### 4.3.3 Perennial organisms: Ricker recruitment without covED

Models with Ricker recruitment have sometimes been proposed without covED, for example by Ellner (1989). The equation for the dynamics of  $N_i(t)$  remain as defined for the previous example (4.10), but  $D$  now gains a subscript, because it differs by species:

$$D_i(t) = \sum_{j=1}^n \alpha_{ij} N_j(t). \quad (4.12)$$

Here the  $\alpha_{ij}$  are positive constants with  $\alpha_{ii} = 1$  and the others less than 1. Thus, the  $\alpha$  coefficients differ by the species  $i$  experiencing the competition, not just by the species  $j$  causing it. However, the most important difference with the previous model is the assumption that  $D_i(t)$  involves the adult densities  $N_j(t)$  not the juvenile densities  $\exp(E_j(t))N_j(t)$ . It is thus assumed that juveniles suffer from competition with the

adult organisms, which restricts their ability to recruit into the adult population, while in the previous model instead, the juveniles are competing with each other for specific resources that juveniles need to make this transition. The outcome mathematically is that covED is not present, although  $\gamma_i$  remains negative. Thus, the storage effect coexistence mechanism cannot function. However, coexistence can occur by other means. As  $S(t)$  no longer plays any additional role, we shall assume  $S(t) \equiv 1$  and that  $((E_i(t))_{i=1,\dots,n})_{t \in \mathbb{Z}^+}$  is a sequence of i.i.d random variables. The next lemma gives conditions under which we can apply our general results.

**Lemma 4.2** *Suppose  $\mathbb{E}[(E_i(t))^2] < \infty$ . Then the model given by (4.12) satisfies Assumption 2.1 by taking a small enough  $\gamma_3 > 0$ , and*

$$V(\mathbf{x}) = \sum_j x_j + 1.$$

Assumption 2.2 holds with

$$\phi(\mathbf{x}) = \delta V(\mathbf{x})$$

for some  $\delta > 0$ . Moreover, if the support of  $\max_j (E_j(t) + \ln \delta_j)$  contains values less than 0 then the boundary is accessible.

**Proof** The proof is very similar to the one of Lemma 4.1 and is therefore omitted.  $\square$

Suppose we have only two species in (4.12). Then, using Lemma 4.2 and following the general approach from Sect. 4.1.2 we need to first look at

$$r_i(\delta_0) = \mathbb{E}[\ln f_i(0, E_i(1))] = \mathbb{E}[\ln(\exp(E_i(1)) + 1 - \delta_i)].$$

Suppose that  $r_i(\delta_0) > 0$ ,  $i = 1, 2$ . Then  $X_1, X_2$  survive on their own and have invariant probability measures  $\mu_1, \mu_2$  on  $(0, \infty)$ . By Sect. 4.1.2 we need to look at

$$r_i(\mu_j) = \int_{(0,\infty)} \mathbb{E} \ln[1 - \delta_i + \exp(E_i(1) - \alpha_{ij}x)] \mu_j(dx).$$

Even though it is hard to compute  $r_i(\mu_j)$  explicitly, one can see (Ellner 1989) that if  $|\delta_1 - \delta_2|$  is sufficiently small then  $r_1(\mu_2) > 0$  and  $r_2(\mu_1) > 0$  and one gets coexistence by the results from Sect. 4.1.2. If however,  $r_i(\mu_j) < 0$  we get the extinction of species  $i$ . Coexistence here depends directly on the fact that  $\alpha_{ij}$  was assumed less than 1, while  $\alpha_{jj} = 1$ . This is coexistence that would occur in the absence of environmental fluctuations, and is just the classical outcome that coexistence occurs if intraspecific competition,  $\alpha_{jj}$ , exceeds interspecific competition  $\alpha_{ij}$ . The presence of  $\delta_i$  in this model means that these  $\alpha$  coefficients are not exactly the classical ones of Chesson (2000b), but when the  $\delta_i$  are equal they become equivalent to them. As mentioned above, coexistence by the storage effect cannot occur in this model because covED is zero even though  $\gamma_i < 0$ . Environmental fluctuations may still have a significant

role in coexistence, however, in cases in which the  $\delta_i$  values differ greatly between species. Then, the coexistence mechanism relatively nonlinearity can occur (Kuang and Chesson 2008).

#### 4.3.4 Annual plants with seed banks and predation

Here instead of perennial organisms, we consider annual plant species. This case has been very important in empirical studies of coexistence in a variable environment. For these species, the growing plant survives for less than a year. It flowers and produces seed once, at the end of its life. The seeds can be eaten by animals, for example rodents and ants (“seed predators”), and any seeds that escape become mixed in the soil as the “seed bank.” Seeds in the seed bank generally have environmentally-dependent germination. This means that in any year, only a fraction of the seeds in the seed bank germinate, and that fraction depends on the specific environmental conditions of that year. A fraction of the seeds that do not germinate generally survive to the next year. The state variable in annual plant models,  $N_i(t)$  for species  $i$ , is generally the number of seeds in the seed bank at the end of the year after all the adults have died, and seed predation has already occurred. The density  $P(t)$  of seed predators is an auxiliary variable. The dynamical equations can now be given as

$$\begin{aligned} N_j(t+1) &= N_j(t) \left( s_j(1 - E_j(t)) + E_j(t)Y_j(t)e^{-A_j(t)-C(t)} \right) \\ P(t+1) &= \sum_{j=1}^n E_j(t)Y_j(t)N_j(t)e^{-C(t)}(1 - e^{-A_j(t)}) + s_p P(t) \\ &= P(t) \left( s_p + \sum_{j=1}^n E_j(t)Y_j(t)N_j(t)e^{-C(t)} \frac{1 - e^{-A_j(t)}}{P(t)} \right) \end{aligned}$$

with

$$C(t) = \sum_{k=1}^n G_k(t)N_k(t).$$

and

$$A_j(t) = a_j(\mathbf{N}(t), \mathbf{E}(t))P(t).$$

Here  $E_j(t) \in (0, 1)$  is the germination fraction,  $A_j(t)$  is mortality due to predation and  $Y_j(t)$ , which is maximum seed yield, is expected to vary from year to year depending on how much rain and nutrients there are. Note that here  $Y_j(t)$  is not an auxiliary variable, but an environmental variable. Instead, the predator,  $P(t)$ , is an auxiliary variable. More details about these models can be found in the work by Kuang and Chesson (2009), Kuang and Chesson (2010) and Chesson and Kuang (2010). Most important, by defining  $D_j(t) = C(t) + A_j(t)$ , this model fits the standard  $G(E, D)$  form (4.6). The critical quantity  $\gamma_j$  is negative, and covED is present. This means that



coexistence by the storage effect can occur in this model. Here, there is a potential for two forms of it, a storage effect due to competition and a storage effect due to predation.

We assume that  $(E_j(t), Y_j(t))_{j=1,\dots,n;t \in \mathbb{N}}$  is a sequence of i.i.d random variables and  $\mathbb{E}[(Y_1(1))^2] < \infty$  and  $a_j$  is bounded below by a positive nonrandom constant. Note that  $\frac{1-e^{-A_j(t)}}{P(t)}$  is continuous even at  $P(t) = 0$ . Due to the fact that  $E_j(t)N_j(t)e^{-C(t)}$  and  $\frac{1-e^{-A_j(t)}}{P(t)}$  are bounded above by a constant as long as  $a_j$  is bounded below, we can follow arguments in Lemma 4.1 to show that our assumptions hold.

#### 4.3.5 Additive models: Unbounded offspring distribution

Theoretical ecologists often use models of the form

$$N_i(t+1) = N_i(t)e^{E_i(t)-f_i(\mathbf{N}(t))} \quad (4.13)$$

where  $f_i$  is a positive function representing the survivorship and  $\mathbb{E}_1(t), \dots, \mathbb{E}_n(t)$  are continuous random variables, and generally unbounded. For example, they could be normally distributed with the nonzero variance representing the  $\ln$  of mean number of offspring produced by the individuals of each species at time  $t$ . We can identify  $f_i(\mathbf{N}(t))$  as  $D_i(t)$  and we see that this model takes the  $G(E, D)$  form of the previous models, but because  $\ln G_i(t) = E_i(t) - D_i(t)$ ,  $\gamma_i = 0$ , and the storage effect cannot occur.

For specific forms of  $f_i$  we can show explicitly that our assumptions hold. Moreover, if  $(E_1(t), \dots, E_n(t))_{t \in \mathbb{N}}$  is i.i.d one can also show for specific  $f_i$ 's that the process is irreducible (see Ellner (1989) in the case when  $f_i$  defines a Ricker model). Note that when the support of  $E_i$  is  $\mathbb{R}$  the dynamics from (4.13) will not live in a compact state space so one cannot use the results from Benaïm and Schreiber (2019).

#### 4.3.6 Additive models: discrete Lotka–Volterra

Consider  $n$  species interacting according to

$$X_i(t+1) = X_i(t) \exp \left( b_i(t) + \sum_j a_{ij}(t) X_j(t) \right), i = 1, 2, \dots, n. \quad (4.14)$$

These equations have been introduced by Hofbauer et al. (1987) as discrete time analogues of Lotka–Volterra differential equations. This form of discrete Lotka–Volterra is what we have called Ricker above. To draw the relationships with these models, we can equate  $E_i(t)$  with  $b_i(t)$  and  $D_i(t)$  with  $-\sum_j a_{ij}(t) X_j(t)$ . Then it is seen that  $\gamma_i = 0$  and the storage effect cannot occur. Moreover, as  $D_i(t)$  is linear, relative non-linearity (Chesson 1994) cannot occur either. In fact, the only role that the variable environment has is to allow the model to explore all of the state space, although this is an important role that leads to more complete results than in the deterministic case.

These equations have further been analyzed by Schreiber (2011) and Benaïm and Schreiber (2019) in a stochastic setting. Suppose the system is hierarchically ordered, that is, there exists a permutation of the indices such that  $a_i^{ii} < 0$  for all  $i, t \in \mathbb{Z}_+$  and  $a^{ij}(t) \leq 0$  for all  $i \leq j$  and  $t \in \mathbb{Z}_+$ . By Hofbauer et al. (1987) and Benaïm and Schreiber (2019), if the system is hierarchically ordered, the coefficients  $a_{ii}(t)$  are bounded above by some negative number, and the coefficients  $a_{ij}(t), b_i(t)$  are bounded, one can show using a comparison argument that there exists  $K > 0$  such that  $\mathbf{X}(t)$  enters  $[0, K]^n$  and never leaves it. We can then work under the assumption of a compact state space.

**Remark 4.4** If the state space of  $\mathbf{X}(t)$  is not compact, the system (4.14) can experience abrupt fluctuations from high densities to very small densities. These sudden crashes make it impossible to use the persistence and extinction criteria we have developed—see Section 5 in the work by Chesson (1982) for a simple example of why violent population declines cannot be allowed.

We can make use of the linearity of the system (4.14) and Proposition 2.1 to compute  $r_i(\mu)$  for any ergodic measure  $\mu$ . In order to showcase the results, assume there are only two species, so that

$$\begin{aligned} X_1(t+1) &= X_1(t)e^{b_1(t+1)+a_{11}(t+1)X_1(t)+a_{12}(t+1)X_2(t)}, \\ X_2(t+1) &= X_2(t)e^{b_2(t+1)+a_{22}(t+1)X_2(t)+a_{21}(t+1)X_1(t)}. \end{aligned} \quad (4.15)$$

We first look at the Dirac delta measure  $\delta_0$  at the origin  $(0, 0)$

$$r_i(\delta_0) = \mathbb{E} \ln b_i(1), i = 1, 2.$$

If  $r_i(\delta_0) > 0$  then species  $i$  survives on its own and converges to a unique invariant probability measure  $\mu_i$  supported on  $\mathcal{S}_+^i := \{\mathbf{x} \in \mathcal{S} \mid x_i \neq 0, x_j = 0, i \neq j\}$ . Moreover,

$$r_i(\mu_i) = \mathbb{E}[b_i(1)] + \mathbb{E}[a_{ii}(1)] \int x_i \mu_i(dx_i) = 0$$

which implies

$$\int x_i \mu_i(dx_i) = \frac{\mathbb{E}[b_i(1)]}{\mathbb{E}[-a_{ii}(1)]}.$$

One can use this to compute the per-capita growth rates

$$r_i(\mu_j) = \mathbb{E}[b_i(1)] + \mathbb{E}[a_{ij}(1)] \int x_j \mu_j(dx_j) = \mathbb{E}[b_i(1)] + \mathbb{E}[a_{ij}(1)] \frac{\mathbb{E}[b_j(1)]}{\mathbb{E}[-a_{jj}(1)]}. \quad (4.16)$$

To ensure that the boundary of the state space is accessible, one can assume for example that the  $a_{ij}$ s and the  $b_i$ 's are absolutely continuous with respect to Lebesgue measure and if  $b_i > 0$  then an interval of the form  $(0, L)$  lies in its support. Having

the expressions (4.16) for  $r_1(\mu_2)$  and  $r_2(\mu_1)$  we can make use of the discussion from Sect. 4.1.2 to classify the dynamics. See Hening (2021) for a more complete discussion of 2d Ricker models.

We note that these results are in a sense more complete than what is known in the deterministic setting for discrete Lotka–Volterra systems where the classification of the long term behavior is not fully known (Ryals and Sacker 2015, 2016; Gyllenberg et al. 2019).

## 5 Applications in continuous time

### 5.1 Structured populations

The survival of an organism is influenced by both biotic (competition for resources, predator–prey interactions) and abiotic (light, precipitation, availability of resources) factors. Since these factors are space-time dependent, all types of organisms have to choose their dispersal strategies: If they disperse they can arrive in locations with different environmental conditions while if they do not disperse they face the temporal fluctuations of the local environmental conditions. The dispersion strategy impacts key attributes of a population including its spatial distribution and temporal fluctuations in its abundance. Continuous-space discrete-time population models that disperse and experience uncorrelated, environmental stochasticity have been studied by Hardin et al. (1988a), Hardin et al. (1988b) and Hardin et al. (1988c). They show that the leading Lyapunov exponent  $r$  of the linearization of the system around the extinction state usually determines the persistence and extinction of the population. Evans et al. (2013) studied a linear stochastic model that describes the dynamics of populations that continuously experience uncertainty in time and space. In Hening et al. (2018) the authors generalized Evans et al. (2013) to a density-dependent model of stochastic population growth that captures the interactions between dispersal and environmental heterogeneity. We will showcase how one can recover and extend the results from Hening et al. (2018).

Suppose we have a population with overlapping generations, which live in a spatio-temporally heterogeneous environment consisting of  $n$  distinct patches. The growth rate of each patch is determined by both deterministic and stochastic environmental inputs. We denote by  $X_i(t)$  the population abundance at time  $t \geq 0$  of the  $i$ th patch and write  $\mathbf{X}(t) = (X_1(t), \dots, X_n(t))$  for the vector of population abundances.

Consider the system

$$dX_i(t) = \left( X_i(t) (a_i - b_i(X_i(t))) + \sum_{j=1}^n D_{ji} X_j(t) \right) dt + X_i(t) dE_i(t), \quad i = 1, \dots, n, \quad (5.1)$$

where  $D_{ij} \geq 0$  for  $j \neq i$  is the per-capita rate at which the population in patch  $i$  disperses to patch  $j$ ,  $D_{ii} = -\sum_{j \neq i} D_{ij}$  is the total per-capita immigration rate out of patch  $i$ ,  $\mathbf{E}(t) = (E_1(t), \dots, E_n(t))^T = \Gamma^T \mathbf{B}(t)$ ,  $\Gamma$  is a  $n \times n$  matrix such that  $\Gamma^T \Gamma = \Sigma = (\sigma_{ij})_{n \times n}$  and  $\mathbf{B}(t) = (B_1(t), \dots, B_n(t))$  is a vector of independent

standard Brownian motions adapted to the filtration  $\{\mathcal{F}_t\}_{t \geq 0}$ . We make the following assumptions.

**Assumption 5.1** For each  $i = 1, \dots, n$  the function  $b_i : \mathbb{R}_+ \mapsto \mathbb{R}$  is locally Lipschitz and vanishing at 0. Furthermore, there are  $M_b > 0$ ,  $\gamma_b > 0$  such that

$$\frac{\sum_{i=1}^n x_i (b_i(x_i) - a_i)}{\sum_{i=1}^n x_i} > \gamma_b \text{ for any } x_i \geq 0, i = 1, \dots, n \text{ satisfying } \sum_{i=1}^n x_i \geq M_b \quad (5.2)$$

**Assumption 5.2** The dispersal matrix  $D$  is *irreducible*.

**Assumption 5.3** The covariance matrix  $\Sigma$  is non-singular.

**Remark 5.1** Condition (5.2) is biologically reasonable because it holds if the  $b_i$ 's are sufficiently large for large  $x_i$ 's. Below, we give some simple scenarios under which Assumption 5.1 is satisfied.

- a) Suppose  $b_i : [0, \infty) \rightarrow [0, \infty)$ ,  $i = 1, \dots, n$  are locally Lipschitz and vanishing at 0. Assume that there exist  $\gamma_b > 0$ ,  $M_b > 0$  such that

$$\inf_{x \in [M_b, \infty)} b_i(x) - a_i - \gamma_b > 0, \quad i = 1, \dots, n$$

Then Assumption 5.1 holds (see Hening et al. 2018).

- b) We note that (a) is satisfied if for  $i = 1, \dots, n$  the function  $b_i : \mathbb{R}_+ \mapsto \mathbb{R}$  is locally Lipschitz, vanishing at 0 and satisfies  $\lim_{x \rightarrow \infty} b_i(x) = \infty$ .
- c) One natural choice for the competition functions, which is widely used throughout the literature, is  $b_i(x) = \kappa_i x$ ,  $x \in (0, \infty)$  for some  $\kappa_i > 0$ . In this case the competition terms become  $-x_i b(x_i) = -\kappa_i x_i^2$ . It is easy to see that these functions satisfy (b) above.

Assumption 5.2 is equivalent to forcing the entries of the matrix  $P_t = \exp(tD)$  to be strictly positive for all  $t > 0$ . This means that it is possible for the population to disperse between any two patches. Assumption 5.3 says that our randomness is non-degenerate, and thus truly  $n$ -dimensional.

We define the total abundance of our population at time  $t \geq 0$  via  $X(t) := \sum_{i=1}^n X_i(t)$  and let  $Y_i(t) := \frac{X_i(t)}{X(t)}$  be the proportion of the total population that is in patch  $i$  at time  $t \geq 0$ . Set  $\mathbf{Y}(t) = (Y_1(t), \dots, Y_n(t))$ . An application of Itô's lemma to (5.1) yields

$$\begin{aligned} dY_i(t) &= Y_i(t) \left( a_i - \sum_{j=1}^n a_j Y_j(t) - b_i(X(t)Y_i(t)) + \sum_{j=1}^n Y_j(t) b_j(X(t)Y_j(t)) \right) dt + \sum_{j=1}^n D_{ji} Y_j(t) dt \\ &\quad + Y_i(t) \left( \sum_{j,k=1}^n \sigma_{kj} Y_k(t) Y_j(t) - \sum_{j=1}^n \sigma_{ij} Y_j(t) \right) dt + Y_i(t) \left[ dE_i(t) - \sum_{j=1}^n Y_j(t) dE_j(t) \right] \\ dX(t) &= X(t) \left( \sum_{i=1}^n (a_i Y_i(t) - Y_i(t) b_i(X(t)Y_i(t))) \right) dt + X(t) \sum_{i=1}^n Y_i(t) dE_i(t) \end{aligned} \quad (5.3)$$

We can rewrite (5.3) in the following compact equation for  $(\mathbf{Y}(t), X(t))$  where  $\mathbf{b}(\mathbf{x}) = (b_1(x_1), \dots, b_n(x_n))$ .

$$\begin{aligned} d\mathbf{Y}(t) &= \left( \text{diag}(\mathbf{Y}(t)) - \mathbf{Y}(t)\mathbf{Y}^\top(t) \right) \Gamma^\top d\mathbf{B}(t) \\ &\quad + \mathbf{D}^\top \mathbf{Y}(t)dt + \left( \text{diag}(\mathbf{Y}(t)) - \mathbf{Y}(t)\mathbf{Y}^\top(t) \right) (\mathbf{a} - \Sigma \mathbf{Y}(t) - \mathbf{b}(X(t)\mathbf{Y}(t)))dt \\ dX(t) &= X(t) [\mathbf{a} - \mathbf{b}(X(t)\mathbf{Y}(t))]^\top \mathbf{Y}(t)dt + X(t)\mathbf{Y}(t)^\top \Gamma^\top d\mathbf{B}(t), \end{aligned} \quad (5.4)$$

where  $\mathbf{Y}(t)$  lies in the simplex  $\Delta := \{(y_1, \dots, y_n) \in \mathbb{R}_+^n : y_1 + \dots + y_n = 1\}$ . Let  $\Delta^\circ = \{(y_1, \dots, y_n) \in \mathbb{R}_+^{n,\circ} : y_1 + \dots + y_n = 1\}$  be the interior of  $\Delta$ .

Consider Eq. (5.4) on the boundary  $((\mathbf{y}, x) : \mathbf{y} \in \Delta, x = 0)$  (that is, we set  $X(t) \equiv 0$  in the equation for  $\mathbf{Y}(t)$ ). We have the following system

$$\begin{aligned} d\tilde{\mathbf{Y}}(t) &= \left( \text{diag}(\tilde{\mathbf{Y}}(t)) - \tilde{\mathbf{Y}}(t)\tilde{\mathbf{Y}}^\top(t) \right) \Gamma^\top d\mathbf{B}(t) \\ &\quad + \mathbf{D}^\top \tilde{\mathbf{Y}}(t)dt + \left( \text{diag}(\tilde{\mathbf{Y}}(t)) - \tilde{\mathbf{Y}}(t)\tilde{\mathbf{Y}}^\top(t) \right) (\mathbf{a} - \Sigma \tilde{\mathbf{Y}}(t))dt \end{aligned} \quad (5.5)$$

on the simplex  $\Delta$ .

Evans et al. (2013) proved that the process  $(\tilde{\mathbf{Y}}(t))_{t \geq 0}$  is an irreducible Markov process, which has the strong Feller property and admits a unique invariant probability measure  $\nu^*$  on  $\Delta$ . Let

$$r_X(\nu^*) = \int_{\Delta} \left( \mathbf{a}^\top \mathbf{y} - \frac{1}{2} \mathbf{y}^\top \Sigma \mathbf{y} \right) \nu^*(d\mathbf{y}). \quad (5.6)$$

**Theorem 5.1** *The following hold:*

- Suppose that  $r_X(\nu^*) > 0$ . The process  $\mathbf{X}(t) = (X_1(t), \dots, X_n(t))_{t \geq 0}$  has a unique invariant probability measure  $\pi$  on  $\mathbb{R}_+^{n,\circ}$  that is absolutely continuous with respect to the Lebesgue measure and

$$\lim_{t \rightarrow \infty} A^t \|P_{\mathbf{X}}(t, \mathbf{x}, \cdot) - \pi(\cdot)\|_{TV} = 0, \quad \mathbf{x} \in \mathbb{R}_+^{n,\circ}, \quad (5.7)$$

for some constant  $A > 0$ . Here  $P_{\mathbf{X}}(t, \mathbf{x}, \cdot)$  is the transition probability of  $(\mathbf{X}(t))_{t \geq 0}$ .

- Suppose that  $r_X(\nu^*) < 0$ . For any  $i = 1, \dots, n$  and any  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}_+^n$ ,

$$\mathbb{P}_{\mathbf{x}} \left\{ \lim_{t \rightarrow \infty} \frac{\ln X_i(t)}{t} = r_X(\nu^*) \right\} = 1. \quad (5.8)$$

## 5.2 Continuous-time models with a resource variable

Suppose we have a guild of  $n$  species  $X_1, \dots, X_n$  whose dynamics is given by

$$dX_i(t) = X_i(t)(c_i(t)R(\mathbf{X}(t)) - m_i) + X_i(t)\sigma_i dB_i(t) \quad (5.9)$$

Here  $R(X_1(t), \dots, X_n(t))$  is the resource abundance and  $c_i(t)$  is the resource uptake rate. The resource abundance is assumed to be given by the algebraic equation

$$R(\mathbf{x}) = R_{\max} - \sum_{j=1}^n c_j(t)x_j(t), \mathbf{x} \in \mathbb{R}_+^n. \quad (5.10)$$

This model is a special case of the resource competition model with fast resource dynamics of Li and Chesson (2016). The  $c_j(t)$  are environmentally varying resource uptake rates and hence positively affect the growth rates of the species while negatively affecting resource growth and abundance. Because resource uptake is assumed fast, (5.10) represents the equilibrium of the resource with the consumer densities at time  $t$ .

One way of modelling the  $c_j(t)$  is by using an Ornstein–Uhlenbeck process

$$d\mathbf{U}(t) = -\gamma(\mathbf{U}(t) - \boldsymbol{\alpha})dt + \mathbf{A}d\mathbf{W}(t) \quad (5.11)$$

where  $\boldsymbol{\alpha}$  is the mean of  $\mathbf{U}(t)$  and  $\mathbf{A}$  is a constant matrix. It is well-known that the process  $\mathbf{U}(t)$  converges as  $t \rightarrow \infty$  to a stationary distribution  $\mu_{\mathbf{X}}$  that is normal and has mean  $\boldsymbol{\alpha}$  and covariance matrix  $M := \mathbf{A}\mathbf{A}^T/(2\gamma)$ . Define

$$c_i(t) = \phi_i(U_i(t)) := \frac{A_i e^{U_i(t)}}{1 + B_i e^{U_i(t)}} + C_i, t \geq 0 \quad (5.12)$$

where  $A_i, C_i, B_i > 0$  then we have  $0 < C_i \leq c_i(t) \leq \frac{A_i}{B_i} + C_i$ .

**Proposition 5.1** Suppose  $(\mathbf{X}(t), \mathbf{U}(t))$  is given by (5.9), (5.10), (5.11) and (5.12). Then, setting  $V(\mathbf{x}, \mathbf{u}) := 1 + |\mathbf{x}| + |\mathbf{u}|^2$ , Assumptions 3.1 and 3.2 will hold for  $(\mathbf{X}(t), \mathbf{U}(t))$ .

**Proof** We can check that there exist constants  $a_1, a_2 > 0$  such that

$$\mathcal{L}|\mathbf{u}|^2 \leq a_1 - a_2\|\mathbf{u}\|^2 \text{ for any } \mathbf{u} \in \mathbb{R}^n. \quad (5.13)$$

Since  $C_i \leq c_i(t) \leq \frac{A_i}{B_i} + C_i$ , we can see that

$$\mathcal{L} \sum_i x_i \leq R_{\max} \max_i \left\{ \frac{A_i}{B_i} + C_i \right\} \sum_i x_i - \min_i \{C_i^2\} \sum_i x_i^2 \quad (5.14)$$

In view of (5.13) and (5.14) and using the Lyapunov function  $V(\mathbf{x}, \mathbf{u}) := 1 + |\mathbf{x}| + |\mathbf{u}|^2$  we can see that Assumptions 3.1 and 3.2 hold for the joint process  $(\mathbf{X}(t), \mathbf{U}(t))$ .  $\square$

Related ecological systems have been studied by Armstrong and McGehee (1980) and Li and Chesson (2016). Most important, this model provides a continuous-time analogue of the models in G(E,D) form in discrete-time examples. In fact, the equations can be rewritten as

$$dX_i(t) = X_i(t)(c_j(t)(R_{\max} - m_i) - c_i(t)D_i(t))dt + X_i(t)\sigma_i dB_i(t)$$

or

$$dX_i(t) = X_i(t)g_i(E_i(t), D_i(t))dt + X_i(t)\sigma_i dB_i(t)$$

where,  $D_i(t) = \sum_{j=1}^n c_j(t)x_j(t)$ , we identify  $c_i(t)$  as  $E_i(t)$ , and  $g_i(E_i(t), D_i(t))$  with  $\ln G_i(E_i(t), D_i(t))$ . We see immediately that  $\gamma_i$  is negative, and covED is present. Of most interest in this case, as shown by Li and Chesson (2016), the  $r_i(\mu)$  can be developed explicitly in terms of  $\gamma_i$  and covED giving a very clear example mathematically of how the storage effect operates. It is the sole mechanism of coexistence in this model.

### 5.3 One species and one auxiliary variable

$$\begin{aligned} dX(t) &= X(t)f(X(t), Y(t))dt + X(t)g(X(t), Y(t))dB(t), \\ dY(t) &= u(Y(t))dt + h(Y(t))dW(t). \end{aligned} \quad (5.15)$$

Suppose we know that  $Y(t)$  has a unique invariant probability measure  $\mu_Y$  on  $\mathbb{R}$  and Assumption 3.1 is satisfied (see Remark 3.2 and Assumption 3.3). Then the long term behaviour is determined by the expected per capita growth rate

$$r_X(\delta_0 \times \mu_Y) = \int_{\mathbb{R}} \left( f(0, y) - \frac{g^2(0, y)}{2} \right) \mu_Y(dy).$$

If  $r_X(\delta_0 \times \mu_Y) > 0$  we have persistence while if  $r_X(\delta_0 \times \mu_Y) < 0$  then  $X$  goes extinct almost surely and

$$\lim_{t \rightarrow \infty} \frac{\ln X(t)}{t} = r_X(\delta_0 \times \mu_Y).$$

### 5.4 Stochastic replicator dynamics

An important class of continuous time dynamics is the one called by Fudenberg and Harris (1992) *stochastic replicator dynamics*. We set  $\Delta := \{\mathbf{x} \in \mathbb{R}_+^n \mid \sum_i x_i = 1\}$  and let  $\Delta_0 = \{\mathbf{x} \in \Delta : x_j = 0 \text{ for some } j\}$  be the extinction set, The fitness of population  $i$  is described by a function  $f_i : \Delta \rightarrow \mathbb{R}$  and the number of individuals in population  $i$  is given by

$$dU_i(t) = U_i(t)(f_i(\mathbf{X}_t) + \sigma_i dB_i(t)), \quad (5.16)$$

where

$$X_i(t) = \frac{U_i(t)}{\sum_j U_j(t)}.$$

The stochastic replicator dynamics has been studied recently by various authors (Imhof 2005; Benaïm et al. 2008; Hofbauer and Imhof 2009; Schreiber 2011). Garay and Hofbauer (2003) studied the permanence and impermanence of deterministic replicator dynamics. Benaïm et al. (2008) looked at these conditions for stochastic replicator dynamics. When the deterministic replicator dynamics is permanent and the noise level small, they showed that the stochastic dynamics admits a unique invariant probability measure whose mass is concentrated near the maximal interior attractor of the unperturbed system. When the deterministic dynamics is impermanent and the noise level small or large, Benaïm et al. (2008) showed that the stochastic dynamics converges to the boundary of the state space at an exponential rate. In Schreiber (2011) the authors were able to give conditions for persistence and convergence to a stationary distribution. Here we recover and strengthen the results of Schreiber (2011) by proving stronger persistence as well as extinction results.

Applying Ito's formula to (5.16) one can see (Schreiber 2011) that

$$dX_i(t) = X_i(t)F_i(\mathbf{X}(t))dt + X_i(t) \sum_j g_{ij}(\mathbf{X}(t))dB_j(t), \quad i = 1, \dots, n$$

where

$$F_i(\mathbf{x}) = f_i(\mathbf{x}) - \sigma_{ii}^2 x_i - \sum_j x_j (f_j(\mathbf{x}) - \sigma_{jj}^2 x_j)$$

and

$$g_{ij}(\mathbf{x}) = (\delta_{ij} - x_j)\sigma_j.$$

It is immediate that for two types the only possible ergodic measures on  $\Delta_0$  are the Dirac masses at  $(1, 0)$  and  $(0, 1)$ . This then yields

$$r_1(\delta_{(0,1)}) = f_1(0, 1) - f_2(0, 1) - \frac{1}{2}(\sigma_1^2 - \sigma_2^2)$$

and

$$r_2(\delta_{(1,0)}) = f_2(1, 0) - f_1(1, 0) - \frac{1}{2}(\sigma_1^2 - \sigma_2^2).$$

We next follow the example studied by Schreiber (2011) and show how we can now give a complete description of the dynamics. Suppose there are three interacting species and  $f_1(\mathbf{x}) = \mu_1 + bx_3$ ,  $f_2(\mathbf{x}) = \mu_2$  and  $f_3(\mathbf{x}) = \mu_3 - cx_1$ . Interactions between types 1 and 3 provide a benefit  $b > 0$  to type 1 and a cost  $c > 0$  to type 3. Assume the following inequality holds

$$\mu_3 - \frac{\sigma_3^2}{2} > \mu_2 - \frac{\sigma_2^2}{2} > \mu_1 - \frac{\sigma_1^2}{2} > 0. \quad (5.17)$$



By Schreiber (2011) it is easy to see that if  $Y_1(0) = 0$ ,  $Y_3(0) > 0$  then  $\mathbf{X}(t)$  converges almost surely to  $(0, 0, 1)$  since  $Y_i(t) = Y_i(0) \exp\left(\left(\mu_i - \frac{\sigma_i^2}{2}\right)t + \sigma_i B_i(t)\right)$  and (5.17) holds. Similarly, if  $Y_3(0) = 0$ ,  $Y_1(0) > 0$  then  $\mathbf{X}(t)$  converges almost surely to  $(0, 1, 0)$ .

One can compute the expected per capita growth rates when  $\mu$  is the Dirac mass function at  $(0, 1, 0)$ ,  $(1, 0, 0)$ , and  $(0, 0, 1)$ :

$$\begin{aligned} r_2(1, 0, 0) &= \mu_2 - \frac{\sigma_2^2}{2} - \mu_1 + \frac{\sigma_1^2}{2} > 0 \\ r_3(1, 0, 0) &= \mu_3 - c - \frac{\sigma_3^2}{2} - \mu_2 + \frac{\sigma_2^2}{2} \\ r_1(0, 1, 0) &= \mu_1 - \frac{\sigma_1^2}{2} - \mu_2 + \frac{\sigma_2^2}{2} < 0 \\ r_3(0, 1, 0) &= \mu_3 - \frac{\sigma_3^2}{2} - \mu_2 + \frac{\sigma_2^2}{2} > 0 \\ r_1(0, 0, 1) &= \mu_1 + b - \frac{\sigma_1^2}{2} - \mu_3 + \frac{\sigma_3^2}{2} \\ r_2(0, 0, 1) &= \mu_2 - \frac{\sigma_2^2}{2} - \mu_3 + \frac{\sigma_3^2}{2} < 0. \end{aligned} \quad (5.18)$$

Note the above imply by Theorem 3.6 or Theorem 3.8 that it is not possible to have ergodic probability measures on  $\Delta_{12} := \{\mathbf{x} \in \Delta \mid x_1 > 0, x_2 > 0, x_3 = 0\}$  or on  $\Delta_{23} := \{\mathbf{x} \in \Delta \mid x_1 = 0, x_2 > 0, x_3 > 0\}$ . We have the following possibilities.

1) Assume

$$b > \mu_3 - \frac{\sigma_3^2}{2} - \mu_1 + \frac{\sigma_1^2}{2} > c. \quad (5.19)$$

Then  $r_1(0, 0, 1) > 0$  and  $r_3(1, 0, 0) > 0$ . This implies that there exists a unique invariant probability measure  $\mu_{13}$  on  $\Delta_{13} := \{\mathbf{x} \in \Delta \mid x_1 > 0, x_2 = 0, x_3 > 0\}$  and by Schreiber (2011) that

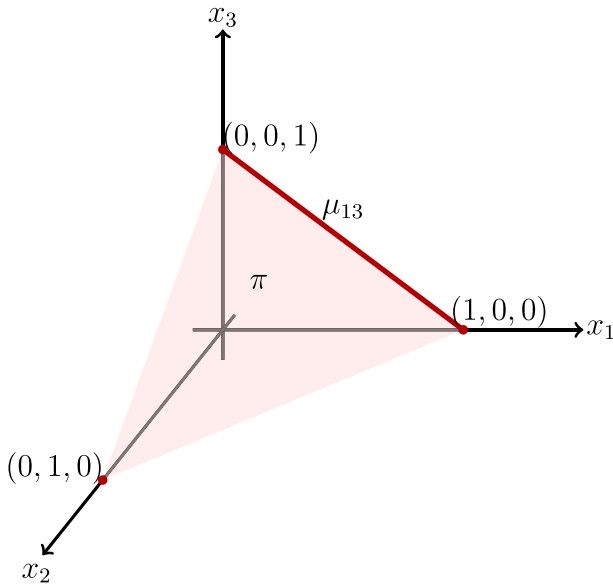
$$r_2(\mu_{13}) = \frac{b\sigma_3^2 - (b-c)\sigma_2^2 - 2br_3 + 2(b-c)r_2 + 2cr_1 + 2bc}{2(b-c)}.$$

If  $r_2(\mu_{13}) > 0$  the populations coexist and  $\mathbf{X}(t)$  converges to its unique stationary distribution  $\pi$  on  $\Delta_+ := \Delta \setminus \Delta_0$ . See Fig. 3 for some intuition of where the different ergodic measures live on the simplex  $\Delta$ .

If  $r_2(\mu_{13}) < 0$  we have

$$\mathbb{P}_{\mathbf{x}} \left\{ \mathcal{U}(\omega) = \{\mu_{13}\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_2(t)}{t} = r_2(\mu_{13}) \right\} = 1$$

for all  $\mathbf{X}(0) = \mathbf{x} \in \Delta_+$ .



**Fig. 3** Coexistence for a three species stochastic replicator system. The red measures signify repellers and  $\pi$  is the invariant probability measure the system converges to

2) Assume (5.19) does not hold and

$$b < \mu_3 - \frac{\sigma_3^2}{2} - \mu_2 + \frac{\sigma_2^2}{2}.$$

Then  $r_1(0, 0, 1) < 0$  and  $r_2(0, 0, 1) < 0$  and  $\delta_{(0,0,1)}$  is a transversal attractor. As a result

$$\mathbb{P}_{\mathbf{x}} \left\{ \mathcal{U}(\omega) = \{\delta_{(0,0,1)}\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_j(t)}{t} = r_j((0, 0, 1)), j = 1, 2 \right\} = 1$$

for all  $\mathbf{X}(0) = \mathbf{x} \in \Delta_+$ .

3) Assume (5.19) does not hold and

$$b, c > \mu_3 - \frac{\sigma_3^2}{2} - \mu_2 + \frac{\sigma_2^2}{2}.$$

The only ergodic invariant measures are the Dirac measures at the vertices.

If  $b > c$  we can show that (Schreiber 2011)

$$\max_i r_i(\mu) > 0, \mu \in \text{Conv}\{\delta_{(1,0,0)}, \delta_{(0,1,0)}, \delta_{(0,0,1)}\}$$

we have coexistence and  $\mathbf{X}(t)$  converges to a unique invariant probability measure  $\pi$  on  $\Delta_+$ .

If  $b < c$  persistence is not possible. However, we are not in the setting of any of our extinction theorems because none of the ergodic measures is an attractor. Different methods have to be used—we show in Hening et al. (2021) that the process converges to the extinction set  $\Delta_0$  exponentially fast.

### 5.5 Lotka Volterra: Two prey and one predator

Assume we have two prey and one predator interacting according to the following Lotka–Volterra system

$$\begin{cases} dX_1(t) = X_1(t)[a_1 - a_{11}X_1(t) - a_{13}X_3(t)]dt + X_1(t)dE_1(t) \\ dX_2(t) = X_2(t)[a_2 - a_{22}X_2(t) - a_{23}X_3(t)]dt + X_2(t)dE_2(t) \\ dX_3(t) = X_3(t)[-a_3 - a_{33}X_3(t) + a_{13}X_1(t) + a_{23}X_2(t)]dt + X_3(t)dE_3(t). \end{cases} \quad (5.20)$$

Assume that  $a_i, b_i, c_{ii} > 0, i = 1, 2, 3, c_{23}, c_{31}, c_{12}, c_{21}, c_{31}, c_{32} \geq 0$ . Assume that  $r_1(\delta_0) = a_1 - \frac{\sigma_{11}}{2} > 0$  and  $r_2(\delta_0) = a_2 - \frac{\sigma_{22}}{2} > 0$  so that species 1 and 2 can survive on their own. This implies by Theorem 3.2 that there exist unique ergodic probability measures  $\mu_1$  and  $\mu_2$  on  $\mathbb{R}_{1+} := (0, \infty) \times \{0\} \times \{0\}$  and  $\mathbb{R}_{2+} := \{0\} \times (0, \infty) \times \{0\}$ .

Note that  $r_3(\delta_0) = -a_3 - \frac{\sigma_{33}}{2} < 0$  which implies that, as expected, the predator cannot survive without any of the prey species.

Next, we can also see that for  $i = 1, 2$

$$0 = r_i(\mu_i) = a_i - \frac{\sigma_{ii}}{2} - b_i \int_{\mathbb{R}_{i+}} x_i \mu(dx_i)$$

which implies

$$\int_{\mathbb{R}_{i+}} x_i \mu(dx_i) = \frac{a_i - \frac{\sigma_{ii}}{2}}{a_{ii}}$$

Using this we can compute for  $j = 1, 2$

$$\lambda_j(\mu_i) = \int_{\partial \mathbb{R}_+^3} \left( a_j + \sum_{\ell=1}^3 a_{j\ell} x_\ell - \frac{\sigma_{jj}}{2} \right) \mu_i(d\mathbf{x}) = a_j + a_{ji} \int_{\partial \mathbb{R}_{i+}^\circ} x_i \mu_i(d\mathbf{x}) - \frac{\sigma_{jj}}{2} = a_j - \frac{\sigma_{jj}}{2} > 0.$$

This implies by Theorem 3.2 that there exists a unique invariant probability measure  $\mu_{12}$  on  $\mathbb{R}_{12+} = (0, \infty) \times (0, \infty) \times \{0\}$ .

The predator's expected per capita growth rates are

$$r_3(\mu_i) = -a_3 - \frac{\sigma_{33}}{2} + a_{i3} \left( \frac{a_i - \frac{\sigma_{ii}}{2}}{a_{ii}} \right), i = 1, 2.$$

Suppose that  $r_3(\mu_i) < 0$ ,  $i = 1, 2$ . Then the faces  $\mathbb{R}_{13+}$ ,  $\mathbb{R}_{23+}$  do not support any invariant probability measures. Lastly, in order to compute  $r_3(\mu_{12})$  we note that

$$r_1(\mu_{12}) = r_2(\mu_{12}) = 0$$

which yields

$$r_3(\mu_{12}) = -a_3 - \frac{\sigma_{33}}{2} + a_{13} \left( \frac{a_1 - \frac{\sigma_{11}}{2}}{a_{11}} \right) + a_{23} \left( \frac{a_2 - \frac{\sigma_{22}}{2}}{a_{22}} \right).$$

If  $r_3(\mu_{12}) > 0$  we have the existence of a unique invariant probability measure  $\mu_{123}$  on  $(0, \infty)^3$  and the coexistence of all three species. If  $r_3(\mu_{12}) < 0$  we have the extinction of the predator and the coexistence of the two prey species

$$\mathbb{P}_{\mathbf{x}} \left\{ \mathcal{U}(\omega) = \{\mu_{12}\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_3(t)}{t} = r_3(\mu_{12}) \right\} = 1, \mathbf{x} \in (0, \infty)^3.$$

We note that Lotka–Volterra SDE models have been studied extensively recently (Hening and Nguyen 2018c, b). A full classification of three-dimensional competitive and predator–prey stochastic Lotka–Volterra systems can be found in recent work by Hening et al. (2021). In the model here, each of the prey species experiences intraspecific competition from the negative term  $a_{ii}X_i(t)$  in its growth rate. They also experience predation through the term  $a_{i3}X_3(t)$ , but as the predator depends on these prey species, increasing in density when the prey are abundant, it is a mediator of apparent competition, i.e. the prey species interact with each other negatively through their effects on the predator. In relation to the previous models,  $D_i(t) = a_{ii}X_i(t) + a_{i3}X_3(t)$  for prey species  $i$ . As the model is additive ( $\gamma_i = 0$ ), the storage effect cannot occur and therefore cannot affect the coexistence of the prey species. Moreover, as the per capita growth rates are linear in the species densities, coexistence by relative nonlinearity (Chesson 1994; Kuang and Chesson 2008) cannot occur either. Although the invasion rates,  $r_i(\mu)$ , do depend on variances, implying an effect of stochastic variation on persistence, these effects are more akin to a parameter change than to a strong effect of stochastic variation because substituting  $a_i$  for  $a_i - \sigma_{ii}/2$  in these formulae reproduces the deterministic invasion rates.

## 6 Discussion

Theoretical models in ecology have long been dominated by deterministic models, not because such models were seen as superior or adequate, but because the theory for stochastic models seemed insuperably difficult. Early stochastic models in ecology focused on Markov jump processes (Kendall 1948; Bartlett 1960), which provided some solutions and useful approximations, but such processes are only stochastic because they have discrete population sizes, and their stochastic elements correspond to independent uncertain events in the lives of individuals, such as mortality and birth. Laws of large numbers in such processes mean that for reasonably large population

sizes a deterministic model will be adequate (Kurtz 1970). Stochastic environmental variation, on the other hand, which is pervasive in nature and has strong effects on population dynamics, remains important no matter how large the population might be (Chesson 1978). However, the dynamical equations in such situations are generally nonlinear and only in limited circumstances yield to analytical solution. Some early approaches used linear approximation (May 1974), and such linear approaches are still being used to approximate population fluctuations in some circumstances (Ripa and Ives 2003). It was not until Turelli (1978, 1981) and Turelli and Gillespie (1980) that the essential nonlinearity of models for multiple interacting species was taken into account. However, these models were Lotka–Volterra models in discrete and continuous time, and were linear additive models, i.e. their per capita growth rates had environmental and density effects combining additively, precluding the storage effect as discussed above, and the density effects were linear precluding coexistence due to relative nonlinearity, as discussed above. In this sense they were disappointing because the stochastic environment added very little to the understanding already available from deterministic models. Soon after, however, stochastic environment models in which the environmental fluctuations have strong effects on species coexistence were developed (Chesson and Warner 1981; Chesson 1982; Abrams 1984). Most important, the strong effects of the stochastic environments in these models resulted directly from nonlinearities introduced by adding realistic biology to simpler models in which a stochastic environment had no such strong effects (Chesson 1994).

These findings that stochastic environmental variation may allow robust species coexistence in situations in which coexistence is precluded in a deterministic environment might have been a great stimulus to the development of stochastic environment models in ecology. However, the difficulty of their analysis meant that few theoretical ecologists followed this path. An important development was quadratic approximations to the invasion growth rates,  $r(\mu)$ , in these models (Chesson 1994, 2019), which led to an understanding of the storage effect coexistence mechanism and relative nonlinearity (Armstrong and McGehee 1980; Abrams and Holt 2002). However, a rigorous understanding of how these invasion rates are able to demonstrate stochastic persistence and coexistence beyond two-species settings was lacking until the developments of Schreiber (2011). Since that time, this understanding has been extended, culminating most recently in Benaïm and Schreiber (2019) in which auxiliary variables were introduced. As discussed here, auxiliary variables  $Y$  allow a substantial increase in the richness of the models that can be analyzed by an understanding of the invasion growth rates of the focal interacting species, the  $X$  and  $N$  variables in our development here. As we have discussed above, they allow populations to have a great deal of structure, they allow the environmental fluctuations to be Markovian rather than merely i.i.d, and they allow dynamic mediators of the interactions between species to be considered. For example, here we consider predators to be an auxiliary variable when studying the coexistence of annual plants subject to seed predation, and in other models, e.g. (Li and Chesson 2016; Chesson 2020) resources that the organisms compete for are auxiliary variables. Remarkably, despite addition of the auxiliary variables, stochastic persistence of the focal species is nevertheless still determined by their invasion growth rates alone.

The major contributions of this work are key extensions of the recent results of Benaïm and Schreiber (2019), and thorough connection of these mathematical developments with the ecological literature. Here we have removed the compactness requirement on the state space. Although it has been argued that population densities are necessarily bounded in nature, it is difficult to identify exactly what those bounds are, and many very useful population models do not put strict bounds on how large a population can be. For example, the simple logistic model in a variable environment leads to a gamma distribution for population size (May 1974). Probability distributions in which unrealistically high populations have extremely low probability make much more sense, and the results here accommodate this situation. Similarly past work in discrete time has placed strict limits on the magnitude of environmental variables, with the justification often being the finite capabilities of organisms. However, the central limit theorem, which leads to many valuable models in science, yields the unbounded normal distribution. Many variables in biology are found to be well approximated by the normal distribution on the log scale. Thus, a variable such as reproductive rate might be well approximated by a lognormal distribution, yet with its infinite tail violates the conditions for establishing persistence in the otherwise extremely valuable work of Benaïm and Schreiber (2019). We have found here, that rather than impose strict bounds on random variables, bounds on an expected value involving the multiplication rates and a Lyapunov-like function are sufficient. In addition, we extend previous work by characterizing the situations where some of the species become extinct, and the remainder converge to invariant measures on the boundary of the full state space. As a consequence we are able to gain a full picture of the long-term outcomes for these stochastic population models.

The second major contribution of this work is to explain how these mathematical developments connect with ecological theory showing where it validates some past work, and sets the stage for exciting new developments in the future. Although we do not have space to give the full details of how species coexistence results from environmental fluctuations in these models, we have sketched the major issues, and shown how the various ecological models fit within a family of models whose structure determines when the storage effect coexistence mechanism will be active, promoting coexistence in a variable environment. We have also touched on the coexistence mechanism relative nonlinearity. Both of these mechanisms naturally arise from the mathematical structure of the model, which is a function of the biological details (Chesson 1994). A concise development of formulae for the invasion rates,  $r(\mu)$ , is given in Chesson (2019) exhibiting explicitly the involvement of the storage effect and relative nonlinearity in these rates. Thus, although in general we have not been able to give explicit formulae for the invasion rates in this article, that information is to be found in the ecological literature in general terms in Chesson (1994) and Chesson (2019), and in more specific terms for the individual models from the citations above where the models are discussed. It needs to be emphasized that it can often be difficult to understand how environmental variation will affect outcomes in ecological models, but an understanding of the structure of the model in the terms discussed here goes a long way to determining when strong effects of environmental variation on species coexistence will occur. A great many models in ecology simply take on linear additive forms, which miss key features of nature of special relevance to how

environmental variation can promote species coexistence. Environmental variation can still have effects in such models by permitting the state variables to explore all of the state space, simplifying the results and thereby allowing a simpler and more complete description of the possible long-term outcomes of the dynamics as arise here for the Ricker and Lotka–Volterra models. Of most importance, however, the stage is set for a great deal of synergy between the ecological applications and the mathematical theory of stochastic persistence. Where the mathematical theory can develop rigorous stochastic persistence and extinction conditions, the ecological theory, exemplified by the discussion of the  $G(E, D)$  form for the discrete time multiplication rates, leads to an understanding of where major effects of environmental variation on stochastic extinction and coexistence are to be expected.

**Acknowledgements** The authors acknowledge generous support from the NSF through the Grants DMS-1853463 for Alexandru Hening, DMS-1853467 for Dang Nguyen, and DEB-1353715 for Peter Chesson.

## Appendix A. Persistence proofs

**Lemma A.1** *For any  $\mathbf{z} \in S$ ,  $t \in \mathbb{N}$  we have the bounds*

$$\mathbb{E}_{\mathbf{z}}(V(\mathbf{Z}(t))) \leq \rho^t V(\mathbf{z}) + C \sum_{s=0}^t \rho^s \leq \rho^t V(\mathbf{z}) + \frac{C}{1-\rho},$$

and

$$\begin{aligned} \mathbb{E}_{\mathbf{z}} h(\mathbf{Z}(t), \xi(t)) &\leq \mathbb{E}_{\mathbf{z}} [V(\mathbf{Z}(t)) h(\mathbf{Z}(t), \xi(t))] \leq \rho \mathbb{E}_{\mathbf{z}} V(\mathbf{Z}(t)) + C \\ &\leq \rho^{t+1} V(\mathbf{z}) + \frac{C}{1-\rho}. \end{aligned}$$

If a function  $\psi$  satisfies

$$\lim_{\mathbf{z} \rightarrow \infty} \frac{\psi(\mathbf{z})}{\mathbb{E} V(\mathbf{z}, \xi(t)) h(\mathbf{z}, \xi(t))} = 0$$

then  $\psi$  is  $\mu$ -integrable for any invariant probability measure  $\mu$ . Moreover, if  $\mu(S_+) = 1$  then  $r_i(\mu) = 0$  for any  $i \in I$ .

**Proof** By Assumption A3) we have  $PV(\mathbf{z}) \leq \rho V(\mathbf{z}) + C$ . Using this and the Markov property yields  $\mathbb{E}_{\mathbf{z}}(V(\mathbf{Z}(t+1))) \leq \rho \mathbb{E}_{\mathbf{z}} V(\mathbf{Z}(t)) + C$ . As a result

$$\mathbb{E}_{\mathbf{z}}(V(\mathbf{Z}(t))) \leq \rho^t V(\mathbf{z}) + C \sum_{s=0}^t \rho^s \leq \rho^t V(\mathbf{z}) + \frac{C}{1-\rho}.$$

On the other hand Assumption A3) and the above imply that

$$\begin{aligned}\mathbb{E}_{\mathbf{z}}[V(\mathbf{Z}(t+1))h(\mathbf{Z}(t), \xi(t))] &\leq \rho \mathbb{E}_{\mathbf{z}}V(\mathbf{Z}(t)) + C \\ &\leq \rho^t V(\mathbf{z}) + \frac{C}{1-\rho}.\end{aligned}\quad (\text{A.1})$$

Since by A3) part i)  $V(\mathbf{z}) \geq 1$  we also get that

$$\mathbb{E}_{\mathbf{z}}h(\mathbf{Z}(t), \xi(t)) \leq \rho^{t+1} V(\mathbf{z}) + \frac{C}{1-\rho}.$$

This implies that if  $\mu$  is an ergodic invariant probability measure then

$$\mathbb{E}_{\mu}h(\mathbf{Z}(t), \xi(t)) = \int \mathbb{E}h(\mathbf{z}, \xi(1))\mu(d\mathbf{z}) \leq \frac{C}{1-\rho}.$$

Since

$$|\log F_i(\mathbf{z}, \xi)|^k \leq c_k \max \left\{ F_i(\mathbf{z}, \xi), \frac{1}{F_i(\mathbf{z}, \xi)} \right\}^{\gamma_3} \leq c_k h(\mathbf{z}, \xi)$$

for some  $c_k > 0$ , we see that for any  $k \in \mathbb{N}$  there exists  $C_k > 0$  such that

$$\int \mathbb{E}|\log F_i(\mathbf{z}, \xi(1))|^k \mu(d\mathbf{z}) \leq C_k.$$

The strong law of large numbers for martingales implies that for  $\mu$  almost every  $\mathbf{z}$  we have

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_0^T (\log F_i(\mathbf{Z}(t+1)) - P \log F_i(\mathbf{Z}(t))) = 0, \text{ when } \mathbf{Z}(0) = \mathbf{z}. \quad (\text{A.2})$$

Having (A.2), we can follow the arguments by Benaïm and Schreiber (2019, Lemma 3 and Proposition 1) to obtain that if  $\mu(\mathcal{S}_+^I) = 1$  then  $r_i(\mu) = 0$  for any  $i \in I$ .

Finally, because of the boundedness (A.1), it is standard to show that if a function  $\psi$  satisfies

$$\lim_{\mathbf{z} \rightarrow \infty} \frac{\psi(\mathbf{z})}{\mathbb{E}V(\mathbf{z}, \xi(t))h(\mathbf{z}, \xi(t))} = 0$$

then  $\psi$  is  $\mu$ -integrable for any invariant probability measure  $\mu$ . See Lemma 3.3 in Hening and Nguyen (2018a) for a similar proof.  $\square$

**Lemma A.2** *There exist  $M, C_2, \gamma_4 > 0, \rho_2 \in (0, 1)$  such that*

$$\mathbb{E}_{\mathbf{z}} \left[ V(\mathbf{Z}(1)) \prod_{i=1}^n X_i^{p_i}(1) \right] \leq (\mathbf{1}_{\{|\mathbf{z}| < M\}}(C_2 - \rho_2) + \rho_2) V(\mathbf{z}) \prod_{i=1}^n x_i^{p_i}, \quad \mathbf{z} \in \mathcal{S}$$

for any  $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{R}^n$  satisfying



$$|\mathbf{p}|_1 := \sum |p_i| \leq \gamma_4. \quad (\text{A.3})$$

**Proof** We note that for  $w_1, \dots, w_n > 0$

$$\begin{aligned} \prod_{i=1}^n w_i^{p_i} &\leq \prod_{i=1}^n \left( w_i^{-1} \vee w_i \right)^{|p_i|} \\ &\leq \left( \max_{i=1, \dots, n} \left\{ w_i^{-1} \vee w_i \right\} \right)^{\sum_i |p_i|} \\ &\leq 1 + \frac{\sum_i |p_i|}{\gamma_3} \left( \max_{i=1, \dots, n} \left\{ w_i^{-1} \vee w_i \right\} \right)^{\gamma_3} \end{aligned}$$

if  $\sum_i |p_i| < \gamma_3$ . The last inequality follows from the inequality  $x^p \leq 1 + px$  for  $x \geq 0$ ,  $p \in (0, 1)$ . Thus,

$$\prod_{i=1}^n F_i^{p_i}(\mathbf{z}, \xi) \leq 1 + \frac{|\mathbf{p}|_1}{\gamma_3} h(\mathbf{z}, \xi).$$

Since  $\lim_{\mathbf{z} \rightarrow \infty} V(\mathbf{z}) = \infty$ , we can select  $\rho_1 \in (\rho, 1)$ ,  $M > 0$  and  $C_1 > 0$  such that

$$\rho V(\mathbf{z}) + C \leq (C_1 \mathbf{1}_{\{|z| < M\}} + \rho_1) V(\mathbf{z}).$$

Let  $\gamma_4 \in (0, \gamma_3)$  be such that  $\left(1 + \frac{\gamma_4}{\gamma_3}\right) \rho_1 =: \rho_2 < 1$ . There exists  $C_2 > 0$  satisfying

$$(C_1 \mathbf{1}_{\{|z| < M\}} + \rho_1) \left(1 + \frac{\gamma_4}{\gamma_3}\right) \leq (C_2 - \rho_2) \mathbf{1}_{\{|z| < M\}} + \rho_2.$$

The above estimates together with (2.1), Lemma A.1, and (A.3) yield

$$\begin{aligned} \mathbb{E}_{\mathbf{z}} \left[ V(\mathbf{Z}(1)) \prod_{i=1}^n X_i^{p_i}(1) \right] &= \prod_{i=1}^n x_i^{p_i} \mathbb{E} \left[ V(\mathbf{x} \circ F(\mathbf{z}, \xi), G(\mathbf{z}, \xi)) \prod_{i=1}^n F_i^{p_i}(\mathbf{z}, \xi) \right] \\ &\leq \prod_{i=1}^n x_i^{p_i} \mathbb{E} \left[ V(\mathbf{x} \circ F(\mathbf{z}, \xi), G(\mathbf{z}, \xi)) \left(1 + \frac{|\mathbf{p}|_1}{\gamma_3} h(\mathbf{z}, \xi) \right) \right] \\ &\leq \prod_{i=1}^n x_i^{p_i} (\rho V(\mathbf{z}) + C) \left(1 + \frac{|\mathbf{p}|_1}{\gamma_3}\right) \\ &\leq \prod_{i=1}^n x_i^{p_i} (C_1 \mathbf{1}_{\{|z| < M\}} + \rho_1) V(\mathbf{z}) \left(1 + \frac{|\mathbf{p}|_1}{\gamma_3}\right) \\ &\leq ((C_2 - \rho) \mathbf{1}_{\{|z| < M\}} + \rho_2) \prod_{i=1}^n x_i^{p_i} V(\mathbf{z}). \end{aligned}$$

□

We will denote by  $B_M := \{\mathbf{u} \in \mathcal{S} : \|\mathbf{u}\| \leq M\}$  the closed ball of radius  $M > 0$  around the origin. Let  $\mathcal{M}$  be the set of ergodic invariant probability measures of  $\mathbf{X}$  supported on the boundary  $\mathcal{S}_0 := \partial\mathbb{R}_+^n \times \mathbb{R}^{k_0}$ . Remember that for a subset  $\widetilde{\mathcal{M}} \subset \mathcal{M}$  we denote by  $\text{Conv}(\widetilde{\mathcal{M}})$  the convex hull of  $\widetilde{\mathcal{M}}$ , that is the set of probability measures  $\pi$  of the form  $\pi(\cdot) = \sum_{\mu \in \widetilde{\mathcal{M}}} p_\mu \mu(\cdot)$  with  $p_\mu \geq 0$ ,  $\sum_{\mu \in \widetilde{\mathcal{M}}} p_\mu = 1$ .

Consider  $\mu \in \mathcal{M}$ . Assume  $\mu(\{0\} \times \mathbb{R}^{k_0}) = 0$ . Since  $\mu$  is ergodic there exist  $0 < n_1 < \dots < n_k \leq n$  such that  $\text{supp}(\mu) \subset S^\mu := \mathbb{R}_+^\mu \times \mathbb{R}^{k_0}$  where

$$\mathbb{R}_+^\mu := \{(x_1, \dots, x_n) \in \mathbb{R}_+^n : x_i = 0 \text{ if } i \in I_\mu^c\}$$

for  $I_\mu := \mathcal{S}(\mu) = \{n_1, \dots, n_k\}$  and  $I_\mu^c := \{1, \dots, n\} \setminus \{n_1, \dots, n_k\}$ .

$$\mathbb{R}_+^{\mu, \circ} := \{(x_1, \dots, x_n) \in \mathbb{R}_+^n : x_i = 0 \text{ if } i \in I_\mu^c \text{ and } x_i > 0 \text{ if } i \in I_\mu\}$$

and  $\partial\mathbb{R}_+^\mu := \mathbb{R}_+^\mu \setminus \mathbb{R}_+^{\mu, \circ}$ .

The following condition ensures persistence.

**Assumption A.1** For any  $\mu \in \text{Conv}(\mathcal{M})$  one has

$$\max_i r_i(\mu) > 0,$$

where

$$r_i(\mu) := \int_{\mathcal{S}_0} [\mathbb{E} \ln F_i(\mathbf{z}, \xi)] \mu(d\mathbf{z}).$$

Remember that the occupation measures are defined as

$$\Pi_{t, \mathbf{z}}(\cdot) = \frac{1}{t} \sum_{s=0}^t \mathbb{P}_{\mathbf{z}}(\mathbf{Z}(s) \in \cdot) ds, \quad \mathbf{z} \in \mathcal{S}, t \in \mathbb{Z}_+$$

**Lemma A.3** Suppose the following

- The sequences  $(\mathbf{z}_k)_{k \in \mathbb{N}} \subset \mathbb{R}_+^n \times \mathbb{R}^{k_0}$ ,  $(T_k)_{k \in \mathbb{N}} \subset \mathbb{N}$  are such that  $\|\mathbf{z}_k\| \leq M$ ,  $T_k > 1$  for all  $k \in \mathbb{N}$  and  $\lim_{k \rightarrow \infty} T_k = \infty$ .
- The sequence  $(\Pi_{T_k, \mathbf{z}_k})_{k \in \mathbb{N}}$  converges weakly to an invariant probability measure  $\pi$ .
- The function  $h : \mathbb{R}_+^n \times \mathbb{R}^{k_0} \rightarrow \mathbb{R}$  is any continuous function satisfying

$$\lim_{\mathbf{z} \rightarrow \infty} \frac{|h(\mathbf{z})|}{V(\mathbf{z}) \max_{i=1}^n \{(x_i^{\gamma_3} \wedge x_i^{-\gamma_3})\}} = 0$$

Then one has

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}_+^n} h(\mathbf{x}) \Pi_{T_k, \mathbf{z}_k}(d\mathbf{x}) = \int_{\mathbb{R}_+^n} h(\mathbf{x}) \pi(d\mathbf{x}).$$

**Proof** The proof is almost identical to that of Lemma 3.4 by Hening and Nguyen (2018a) and is therefore omitted.  $\square$

It is shown in Schreiber (2011, Lemma 4) by the min–max principle that Assumption A.1 is equivalent to the existence of  $\mathbf{p} > 0$  such that

$$\min_{\mu \in \mathcal{M}} \left\{ \sum_i p_i r_i(\mu) \right\} := 2r^* > 0. \quad (\text{A.4})$$

By rescaling if necessary, we can assume that  $|\mathbf{p}|_1 = \gamma_4$ .

**Lemma A.4** Suppose that Assumption A.1 holds. Let  $\mathbf{p}$  and  $r^*$  be as in (A.4). There exists an integer  $T^* > 0$  such that, for any  $T > T^*$ ,  $\mathbf{x} \in \partial \mathbb{R}_+^n$ ,  $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in B_M$  one has

$$\sum_{t=0}^T \mathbb{E}_{\mathbf{z}} \left( \ln V(\mathbf{Z}(t+1)) - \ln V(\mathbf{Z}(t)) - \sum p_i \ln F_i(\mathbf{Z}(t), \xi(t)) \right) \leq -r^*(T+1). \quad (\text{A.5})$$

**Proof** In view of Lemma A.1,

$$\sup_{t \in \mathbb{N}, \|\mathbf{z}\| \leq M} \mathbb{E}_{\mathbf{z}} V(\mathbf{Z}(t)) < \infty$$

which implies, since

$$\frac{\mathbb{E}_{\mathbf{z}} \ln V(\mathbf{Z}(T))}{T+1} \leq \frac{\mathbb{E}_{\mathbf{z}} V(\mathbf{Z}(T))}{T+1}$$

that

$$\lim_{T \rightarrow \infty} \sup_{\|\mathbf{z}\| \leq M} \frac{1}{T+1} \sum_{t=0}^T \mathbb{E}_{\mathbf{z}} (\ln V(\mathbf{Z}(t+1)) - \ln V(\mathbf{Z}(t))) = 0.$$

With (A.4) and Lemma A.2 and Lemma A.3, we can argue by contradiction in the same manner as in Lemma 4.1 by Hening and Nguyen (2018a) to show that

$$\limsup_{T \rightarrow \infty} \frac{1}{T+1} \sup_{\|\mathbf{z}\| \leq M} \sum_{t=0}^T \mathbb{E}_{\mathbf{z}} \left( - \sum p_i \ln F_i(\mathbf{Z}(t), \xi(t)) \right) \leq -2r^*.$$

Combining the two above limits finishes the proof.  $\square$

Let  $n^* \in \mathbb{N}$  be such that

$$\rho_2^{1-n^*} > C_2. \quad (\text{A.6})$$

**Proposition A.1** Define  $U : \mathcal{S}_+ \rightarrow \mathbb{R}_+$  by

$$U(\mathbf{z}) = V(\mathbf{z}) \prod_{i=1}^n x_i^{-p_i}$$

with  $\mathbf{p}$  and  $r^*$  satisfying (A.4) and  $T^* > 0$  satisfying the assumptions of Lemma A.4. There exist numbers  $\theta \in (0, \frac{\gamma_4}{2})$ ,  $K_\theta > 0$ , such that for any  $T \in [T^*, n^* T^*] \cap \mathbb{Z}$  and  $\mathbf{z} \in \mathcal{S}_+$ ,  $\|\mathbf{z}\| \leq M$ ,

$$\mathbb{E}_{\mathbf{z}} U^\theta(\mathbf{Z}(T)) \leq U^\theta(\mathbf{x}) \exp\left(-\frac{1}{2}\theta r^* T\right) + K_\theta.$$

**Proof**

$$\begin{aligned} \ln U(\mathbf{Z}(T)) &= \ln U(\mathbf{Z}(0)) + \sum_{t=0}^{T-1} (\ln U(\mathbf{Z}(t+1)) - \ln U(\mathbf{Z}(t))) \\ &= \ln U(\mathbf{Z}(0)) + G(T) \end{aligned} \quad (\text{A.7})$$

where

$$G(T) = \sum_{t=0}^{T-1} \left( \ln V(\mathbf{Z}(t+1)) - \ln V(\mathbf{Z}(t)) - \sum p_i \ln F_i(\mathbf{Z}(t), \xi(t)) \right). \quad (\text{A.8})$$

In view of (A.7) and Lemma A.2

$$\mathbb{E}_{\mathbf{z}} \exp(G(T)) = \frac{\mathbb{E}_{\mathbf{z}} U(\mathbf{Z}(T))}{U(\mathbf{z})} \leq (C_2)^T. \quad (\text{A.9})$$

Let  $\widehat{U}(\cdot) : \mathbb{R}_+^{n,\circ} \times \mathbb{R}^{\kappa_0} \mapsto \mathbb{R}_+$  be defined by  $\widehat{U}(\mathbf{z}) = V(\mathbf{z}) \prod_{i=1}^n x_i^{p_i}$ . We also have

$$\frac{\mathbb{E}_{\mathbf{z}} \widehat{U}(\mathbf{Z}(T))}{\widehat{U}(\mathbf{z})} \leq (C_2)^T. \quad (\text{A.10})$$

Note that

$$U^{-1}(\mathbf{z}) = \widehat{U}(\mathbf{z}) \frac{1}{V^2(\mathbf{z})} \leq \widehat{U}(\mathbf{z}). \quad (\text{A.11})$$

Using (A.11) and (A.10) yields

$$\begin{aligned}\mathbb{E}_{\mathbf{z}} \exp(-G(T)) &= \frac{\mathbb{E}_{\mathbf{z}} U^{-1}(\mathbf{Z}(T))}{U^{-1}(\mathbf{z})} \\ &\leq \frac{\mathbb{E}_{\mathbf{z}} \widehat{U}(\mathbf{Z}(T))}{V^2(\mathbf{z}) U^{-1}(\mathbf{z})} \\ &\leq \frac{\mathbb{E}_{\mathbf{z}} \widehat{U}(\mathbf{Z}(T))}{\widehat{U}(\mathbf{z})} \\ &\leq (C_2)^T.\end{aligned}\tag{A.12}$$

By (A.9) and (A.12) the assumptions of Hening and Nguyen (2018a, Lemma 3.5) hold for the random variable  $G(T)$ . Therefore, there exists  $\tilde{K}_2 \geq 0$  such that

$$0 \leq \frac{d^2 \tilde{\phi}_{\mathbf{z}, T}}{d\theta^2}(\theta) \leq \tilde{K}_2 \text{ for all } \theta \in \left[0, \frac{1}{2}\right), \mathbf{z} \in \mathbb{R}_+^{n, \circ} \times \mathbb{R}^{K_0}, \|\mathbf{z}\| \leq M, T \in [T^*, n^* T^*] \cap \mathbb{Z}$$

where

$$\tilde{\phi}_{\mathbf{z}, T}(\theta) = \ln \mathbb{E}_{\mathbf{z}} \exp(\theta G(T)).$$

In view of Lemma A.4 and the Feller property of  $(\mathbf{Z}(t))$ , there exists a  $\tilde{\delta} > 0$  such that if  $\|\mathbf{z}\| \leq M$ ,  $\text{dist}(\mathbf{z}, \partial \mathbb{R}_+^n) < \tilde{\delta}$  and  $T \in [T^*, n^* T^*] \cap \mathbb{Z}$  then

$$\mathbb{E}_{\mathbf{z}} G(T) \leq -\frac{3}{4} r^* T.\tag{A.13}$$

Another application of Hening and Nguyen (2018a, Lemma 3.5) yields

$$\begin{aligned}\frac{d \tilde{\phi}_{\mathbf{z}, T}}{d\theta}(0) &= \mathbb{E}_{\mathbf{z}} G(T) \leq -\frac{3}{4} r^* T \text{ for } \mathbf{z} \in \mathbb{R}_+^{n, \circ} \times \mathbb{R}^{K_0}, \|\mathbf{z}\| \leq M, \text{dist}(\mathbf{z}, \partial \mathbb{R}_+^n) \\ &< \tilde{\delta}, T \in [T^*, n^* T^*] \cap \mathbb{Z}.\end{aligned}$$

By a Taylor expansion around  $\theta = 0$ , for  $\|\mathbf{z}\| \leq M$ ,  $\text{dist}(\mathbf{z}, \partial \mathbb{R}_+^n) < \tilde{\delta}$ ,  $T \in [T^*, n^* T^*] \cap \mathbb{Z}$  and  $\theta \in [0, \frac{1}{2})$  we have

$$\tilde{\phi}_{\mathbf{z}, T}(\theta) \leq -\frac{3}{4} r^* T \theta + \theta^2 \tilde{K}_2.$$

If we choose any  $\theta \in (0, \frac{1}{2})$  satisfying  $\theta < \frac{r^* T^*}{4 \tilde{K}_2}$ , we obtain that

$$\tilde{\phi}_{\mathbf{z}, T}(\theta) \leq -\frac{1}{2} r^* T \theta \text{ for all } \mathbf{z} \in \mathbb{R}_+^{n, \circ} \times \mathbb{R}^{K_0}, \|\mathbf{z}\| \leq M, \text{dist}(\mathbf{z}, \partial \mathbb{R}_+^n) < \tilde{\delta}, T \in [T^*, n^* T^*] \cap \mathbb{Z}.\tag{A.14}$$

In light of (A.14), we have for all  $\theta < \frac{r^*T^*}{4\tilde{K}_2}$ ,  $\|\mathbf{z}\| \leq M$ ,  $0 < \text{dist}(\mathbf{z}, \partial\mathbb{R}_+^n) < \tilde{\delta}$ ,  $T \in [T^*, n^*T^*]$  that

$$\frac{\mathbb{E}_{\mathbf{z}} U^\theta(\mathbf{Z}(T))}{U^\theta(\mathbf{z})} = \exp \tilde{\phi}_{\mathbf{z}, T}(\theta) \leq \exp\left(-\frac{1}{2}r^*T\theta\right). \quad (\text{A.15})$$

In view of Lemma A.2, we have for  $\mathbf{z}$  satisfying  $\|\mathbf{z}\| \leq M$ ,  $\text{dist}(\mathbf{z}, \partial\mathbb{R}_+^n) \geq \tilde{\delta}$  and  $T \in [T^*, n^*T^*]$  that

$$\mathbb{E}_{\mathbf{z}} U^\theta(\mathbf{Z}(T)) \leq (C_2)^{\theta n^*T^*} \sup_{\|\mathbf{z}\| \leq M, \text{dist}(\mathbf{z}, \partial\mathbb{R}_+^n) \geq \tilde{\delta}} \{U^\theta(\mathbf{z})\} =: K_\theta < \infty. \quad (\text{A.16})$$

Combining (A.15) and (A.16) we are done.  $\square$

**Theorem A.1** Suppose that Assumption A.1 holds. Let  $\theta$  be as in Proposition A.1,  $T^*$  as in Lemma A.4 and  $n^*$  as in (A.6). There exist numbers  $\kappa = \kappa(\theta, T^*) \in (0, 1)$  and  $\tilde{K} = \tilde{K}(\theta, T^*) > 0$  such that

$$\mathbb{E}_{\mathbf{z}} U^\theta(\mathbf{Z}(n^*T^*)) \leq \kappa U^\theta(\mathbf{z}) + \tilde{K} \text{ for all } \mathbf{z} \in \mathbb{R}_+^{n, \circ} \times \mathbb{R}^{k_0}. \quad (\text{A.17})$$

We have

$$\limsup_{t \rightarrow \infty} \mathbb{P}_{\mathbf{z}}\{|X_i(t)| \vee |X_i^{-1}(t)| > m \text{ for some } i = 1, \dots, n\} \leq c_2 m^{-c_3} \quad (\text{A.18})$$

for some positive  $c_2, c_3 > 0$ . Moreover, for any compact set  $K \subset \mathbb{R}_+^{n, \circ} \times \mathbb{R}^{k_0}$ ,

$$\mathbb{P}_{\mathbf{z}}(\tau_K > k) \leq c_K U^\theta(\mathbf{z}) \kappa^k \quad (\text{A.19})$$

If the Markov chain  $\mathbf{Z}(t)$  is irreducible and aperiodic on  $\mathbb{R}_+^{n, \circ} \times \mathbb{R}^{k_0}$ , and a compact set is petite, then there is  $c_4 > 1$  such that

$$c_4^t \|P_t(\mathbf{z}, \dots) - \pi\|_{TV} \rightarrow 0 \text{ as } t \rightarrow \infty.$$

**Proof** Define

$$\tau = \inf\{t \geq 0 : \|\mathbf{Z}(t)\| \leq M\}. \quad (\text{A.20})$$

By Lemma A.2 for all  $\mathbf{z} \in S$

$$PU(\mathbf{z}) \leq \rho_2 U(\mathbf{z}), \quad \|\mathbf{z}\| \geq M.$$

This implies that the process  $\rho_2^{-t} U(\mathbf{Z}(t))$  is a supermartingale and therefore

$$\mathbb{E}_{\mathbf{z}} \left[ \rho_2^{-\theta(\tau \wedge n^*T^*)} U^\theta(\mathbf{Z}(\tau \wedge n^*T^*)) \right] \leq U^\theta(\mathbf{z}), \quad \mathbf{z} \in S.$$

Thus,

$$\begin{aligned}
 U^\theta(\mathbf{Z}) &\geq \mathbb{E}_{\mathbf{Z}} \left[ \rho_2^{-\theta(\tau \wedge n^* T^*)} U^\theta(\mathbf{Z}(\tau \wedge n^* T^*)) \right] \\
 &= \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \leq (n^*-1)T^*\}} \rho_2^{-\theta(\tau \wedge n^* T^*)} U^\theta(\mathbf{Z}(\tau \wedge n^* T^*)) \right] \\
 &\quad + \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T^* < \tau < n^* T^*\}} \rho_2^{-\theta(\tau \wedge n^* T^*)} U^\theta(\mathbf{Z}(\tau \wedge n^* T^*)) \right] \\
 &\quad + \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \geq n^* T^*\}} \rho_2^{-\theta(\tau \wedge n^* T^*)} U^\theta(\mathbf{Z}(\tau \wedge n^* T^*)) \right] \quad (\text{A.21}) \\
 &\geq \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \leq (n^*-1)T^*\}} U^\theta(\mathbf{Z}(\tau)) \right] \\
 &\quad + \rho_2^{-\theta(n^*-1)T^*} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T^* < \tau < n^* T^*\}} U^\theta(\mathbf{Z}(\tau)) \right] \\
 &\quad + \rho_2^{-\theta n^* T^*} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \geq n^* T^*\}} U^\theta(\mathbf{Z}(n^* T^*)) \right].
 \end{aligned}$$

By the strong Markov property of  $\mathbf{Z}(t)$  and Proposition A.1, we obtain

$$\begin{aligned}
 &\mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \leq (n^*-1)T^*\}} U^\theta(\mathbf{Z}(n^* T^*)) \right] \\
 &\leq \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \leq (n^*-1)T^*\}} \left[ K_\theta + e^{-\frac{1}{2}\theta r^*(n^* T^* - \tau)} U^\theta(\mathbf{Z}(\tau)) \right] \right] \quad (\text{A.22}) \\
 &\leq K_\theta + \exp \left( -\frac{1}{2}\theta r^* T^* \right) \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \leq (n^*-1)T^*\}} U^\theta(\mathbf{Z}(\tau)) \right]
 \end{aligned}$$

Similarly, the strong Markov property of  $\mathbf{Z}(t)$ , Jensen's inequality and Lemma A.2 imply

$$\begin{aligned}
 &\mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T^* < \tau < n^* T^*\}} U^\theta(\mathbf{Z}(n^* T^*)) \right] \\
 &\leq \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T^* < \tau < n^* T^*\}} C_2^{\theta(n^* T^* - \tau)} U^\theta(\mathbf{Z}(\tau)) \right] \quad (\text{A.23}) \\
 &\leq C_2^{\theta T^*} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T^* < \tau < n^* T^*\}} U^\theta(\mathbf{Z}(\tau)) \right].
 \end{aligned}$$

Applying (A.22) and (A.23) to (A.21) yields

$$\begin{aligned}
 U^\theta(x) &\geq \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \leq (n^*-1)T^*\}} U^\theta(\mathbf{Z}(\tau)) \right] \\
 &\quad + \rho_2^{-\theta(n^*-1)T^*} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T^* < \tau < n^* T^*\}} U^\theta(\mathbf{Z}(\tau)) \right] \\
 &\quad + \rho_2^{-\theta n^* T^*} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \geq n^* T^*\}} U^\theta(\mathbf{Z}(n^* T^*)) \right] \\
 &\geq \exp \left( \frac{1}{2}\theta r^* T^* \right) \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \leq (n^*-1)T^*\}} U^\theta(\mathbf{Z}(n^* T^*)) \right] - \exp \left( \frac{1}{2}\theta r^* T^* \right) K_\theta \quad (\text{A.24}) \\
 &\quad + C_2^{-\theta T^*} \rho_2^{-\theta(n^*-1)T^*} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T^* < \tau < n^* T^*\}} U^\theta(\mathbf{Z}(n^* T^*)) \right] \\
 &\quad + \rho_2^{-\theta n^* T^*} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \geq n^* T^*\}} U^\theta(\mathbf{Z}(n^* T^*)) \right] \\
 &\geq \kappa^{-1} \mathbb{E}_{\mathbf{Z}} U^\theta(\mathbf{Z}(n^* T^*)) - K_\theta \exp \left( \frac{1}{2}\theta \rho^* T^* \right)
 \end{aligned}$$

where  $\kappa = \max \left\{ \exp \left( -\frac{1}{2} \theta r^* T^* \right), C^{\theta T^*} \rho_2^{\theta(n^*-1)T^*}, \rho_2^{\theta n^* T^*} \right\} < 1$  by (A.6). The proof of (A.17) is complete by taking

$$\tilde{K} = K_\theta \exp \left( \frac{1}{2} \theta \rho^* T^* \right) \kappa.$$

Having (A.17), the claims (A.19) and (A.18) follow by Benaïm and Schreiber (2019, Proposition 3.3 and Theorem 3.1).  $\square$

## Appendix B. Extinction Proofs

For  $I \subset \{1, \dots, n\}$ , denote by  $\mathcal{M}^I, \mathcal{M}^{I,+}, \mathcal{M}^{I,\partial}$  the set of ergodic probability measures on  $\mathcal{S}^I, \mathcal{S}_+^I, \mathcal{S}_0^I$  respectively.

**Lemma B.1** *Assume that there exists a function  $\phi : \mathcal{S} \rightarrow \mathbb{R}_+$  and constants  $C, \delta_\phi > 0$  such that for all  $\mathbf{z} \in \mathcal{S}$*

$$PV(\mathbf{z}) \leq V(\mathbf{z}) - \phi(\mathbf{z}) + C \quad (\text{B.1})$$

and

$$\mathbb{E}_{\mathbf{z}} (V(\mathbf{Z}(1)) - PV(\mathbf{z}))^2 + \mathbb{E} |\log F(\mathbf{z}, \xi(1)) - \mathbb{E} \log F(\mathbf{z}, \xi(1))|^2 \leq \delta_\phi \phi(\mathbf{z}). \quad (\text{B.2})$$

Then, the family of random occupation measures  $(\tilde{\Pi}_t)_{t \in \mathbb{N}}$  is tight. and with probability one

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T (\log F(\mathbf{Z}(t), \xi(t)) - \mathbb{E} [\log F(\mathbf{Z}(t), \xi(t)) | \mathcal{F}_t]) = 0. \quad (\text{B.3})$$

where  $(\mathcal{F}_t)_{t \in \mathbb{N}}$  is the filtration generated by the process  $\mathbf{Z}$ .

**Proof** Suppose  $\mathbf{Z}(0) = \mathbf{z} \in \mathcal{S}$ . We have

$$\begin{aligned} V(\mathbf{Z}(t+1)) &\leq V(\mathbf{Z}(t)) - \phi(\mathbf{Z}(t)) + C + (V(\mathbf{Z}(t+1)) - PV(\mathbf{Z}(t))) \\ &\leq V(\mathbf{Z}(t)) - \frac{1}{2} \phi(\mathbf{Z}(t)) + 2C - \left( \frac{1}{2} \phi(\mathbf{Z}(t)) + C - (V(\mathbf{Z}(t+1)) - PV(\mathbf{Z}(t))) \right) \end{aligned} \quad (\text{B.4})$$

We see from (B.2) that the quadratic variation of the martingale  $V(\mathbf{Z}(t+1)) - PV(\mathbf{Z}(t))$  is bounded by  $\delta_\phi(\mathbf{Z}(t))$ . As a result we can use the strong law of large numbers for martingales and the bound (B.2), to get

$$\lim_{T \rightarrow \infty} \frac{\frac{1}{T} \sum_{t=0}^T (V(\mathbf{Z}(t+1)) - PV(\mathbf{Z}(t)))}{\frac{1}{T} \sum_{t=0}^T (\phi(\mathbf{Z}(t)) + C)} = \lim_{T \rightarrow \infty} \frac{\sum_{t=0}^T (V(\mathbf{Z}(t+1)) - PV(\mathbf{Z}(t)))}{\sum_{t=0}^T (\phi(\mathbf{Z}(t)) + C)} = 0 \text{ a.s.}$$



which implies

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \left( \frac{1}{2} \phi(\mathbf{Z}(t)) + C - (V(\mathbf{Z}(t+1)) - PV(\mathbf{Z}(t))) \right) \leq 0 \text{ a.s.} \quad (\text{B.5})$$

Taking sums in (B.4), noting that

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_0^T (V(\mathbf{Z}(t+1)) - V(\mathbf{Z}(t))) = \liminf_{T \rightarrow 0} \frac{1}{T} (V(\mathbf{Z}(T+1)) - V(\mathbf{z})) \geq 0 \text{ a.s.}$$

(because  $V$  is nonnegative) and using (B.5) yields that with probability one

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_0^T \left( -\frac{1}{2} \phi(\mathbf{Z}(t)) + 2C \right) \geq 0.$$

As a result

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_0^T \phi(\mathbf{Z}(t)) \leq 4C, \quad (\text{B.6})$$

almost surely. Since  $\lim_{|\mathbf{z}| \rightarrow \infty} \phi(\mathbf{z}) = \infty$ , the boundedness of  $\tilde{\Pi}_T \phi$  in (B.6) implies that the family of randomized occupation measures  $\{\tilde{\Pi}_t, t \in \mathbb{N}\}$  is tight. Moreover, the strong law of large numbers together with (B.6) and (B.2) implies (B.3).  $\square$

**Theorem B.1** *If  $E_1$  is nonempty, then for any  $I \in E_1$ , there exists  $\alpha_I > 0$  such that, for any a compact set  $\mathcal{K}^I \subset \mathcal{S}_+^I$ , we have*

$$\lim_{\text{dist}(\mathbf{z}, \mathcal{K}^I) \rightarrow 0, \mathbf{z} \in \mathcal{S}^0} \mathbb{P}_{\mathbf{z}} \left\{ \lim_{t \rightarrow \infty} \frac{\ln X_i(t)}{t} \leq -\alpha_I, i \in I^c \right\} = 1.$$

**Theorem B.2** *If  $E_2$  is empty or  $\max_i \{r_i(v)\} > 0$  for any  $v$  with  $\mu(\mathcal{S}_+^J) = 1$  for some  $J \in E_2$  and  $\cup_{I \in E_1} \mathcal{S}_+^I$  is accessible then*

$$\sum_{I \in E_1} p_{\mathbf{z}, I} = 1$$

where

$$p_{\mathbf{z}, I} = \mathbb{P}_{\mathbf{z}} \left\{ \emptyset \neq \mathcal{U}(\omega) \subset \text{Conv}\{\mathcal{M}^{I,+}\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_j(t)}{t} \in \{r_j(\mu) : \mu \in \text{Conv}(\mathcal{M}^{I,+})\}, \right. \\ \left. j \in I^c \right\}.$$

**Proof** Once Theorem B.1 is proved, Theorem B.2 can be obtained using the fact that any weak limit of a family of random occupation measures is an invariant probability measure supported on  $S_0$  (Benaim 2018; Hening and Nguyen 2018a) and by using the arguments from Lemma 5.8, Lemma 5.9 and Theorem 5.2 by Hening and Nguyen (2018a).  $\square$

Fix  $I \in E_1$ . Since by Lemma B.1 the family  $(\tilde{\Pi}_t)_{t \in \mathbb{N}}$  of random occupation measures is tight, condition (2.9) is equivalent to the existence of  $0 < \hat{p}_i < \gamma_3/n$ ,  $i \in I$

$$\inf_{v \in \text{Conv}(\mathcal{M}^{I, \partial})} \sum_{i \in I} \hat{p}_i r_i(v) > 0.$$

As a result, there exists a small  $\check{p} \in (0, \gamma_3/n)$  such that

$$\sum_{i \in I} \hat{p}_i r_i(v) - \check{p} \max_{i \notin I} \{r_i(v)\} > 0 \text{ for any } v \in \text{Conv}(\mathcal{M}^{I, \partial}). \quad (\text{B.7})$$

Define  $\tilde{p}_i$  by  $\tilde{p}_i = \hat{p}_i$  if  $i \in I_\mu$  and  $\tilde{p}_i = -\check{p}$  if  $i \in I_\mu^c$ . In view of (B.7), (2.8) and Lemma B.1, there is  $r_e > 0$  such that for any  $v \in \text{Conv}(\mathcal{M}^I)$ ,

$$\sum_{i \in I} \hat{p}_i r_i(v) - \check{p} \max_{i \in I^c} \{r_i(v)\} > 3r_e. \quad (\text{B.8})$$

**Lemma B.2** Let  $I \in E_1$  and suppose that Assumption 2.1 holds. Suppose  $\hat{p}_i$ ,  $\check{p}$ ,  $r_e$  are the quantities from (B.8) and  $n^*$  is defined by (A.6). There exist constants  $T_e \geq 0$ ,  $\delta_e > 0$  such that, for any  $T \in [T_e, n^*T_e] \cap \mathbb{Z}$ ,  $\|\mathbf{z}\| \leq M$ ,  $x_i < \delta_e$ ,  $i \in I^c$ , we have

$$\sum_{t=0}^T \mathbb{E}_{\mathbf{z}} \left( \ln V(\mathbf{Z}(t+1)) - \ln V(\mathbf{Z}(t)) - \sum_{i \in I} \ln F_i(\mathbf{Z}(t), \xi(t)) + \check{p} \max_{i \in I^c} \ln F_i(\mathbf{Z}(t), \xi(t)) \right) dt \leq -r_e(T+1). \quad (\text{B.9})$$

**Proof** This is very similar to the proof of Lemma A.4 and is therefore omitted.  $\square$

**Proposition B.1** Let  $I \in E_1$  and suppose that Assumption 2.1 holds. There exists  $\theta \in (0, 1)$  such that for any  $T \in [T_e, n^*T_e] \cap \mathbb{Z}$  and  $\mathbf{z} \in \mathcal{S}_+$  satisfying  $\|\mathbf{z}\| \leq M$ ,  $x_i < \delta_e$ ,  $i \in I^c$  one has

$$\mathbb{E}_{\mathbf{z}} W_\theta(\mathbf{Z}(T)) \leq \exp \left( -\frac{1}{2} \theta r_e T \right) W_\theta(\mathbf{z})$$

where  $M$ ,  $T_e$ ,  $\hat{p}_i$ ,  $\check{p}$ ,  $\delta_e$ ,  $n^*$  are as in Lemma B.2 and

$$W_\theta(\mathbf{z}) := \sum_{i \in I^c} \left[ V(\mathbf{z}) \frac{x_i^{\check{p}}}{\prod_{j \in I} x_j^{\hat{p}_j}} \right]^\theta, \quad \mathbf{z} \in \mathcal{S}_+.$$

**Proof** For  $i \in I^c$ , let  $W(\mathbf{z}, i) := V(\mathbf{z}) \frac{x_i^{\check{p}}}{\prod_{j \in I} x_j^{\check{p}_j}}$ . Similarly to Proposition A.1, by making use of Lemma B.2, one can find a  $\theta > 0$  such that for  $T \in [T_e, n^* T_e] \cap \mathbb{Z}$ ,  $\mathbf{z} \in \mathbb{R}_+^{n, \circ} \times \mathbb{R}^{k_0}$  with  $\|\mathbf{z}\| \leq M$ , and  $x_i < \delta_e$  we have

$$\mathbb{E} W^\theta(\mathbf{Z}(T), i) \leq \exp\left(-\frac{1}{2} \theta r_e T\right) W^\theta(\mathbf{z}, i).$$

The proof is complete by noting that

$$W_\theta(\mathbf{z}) = \sum_{i \in I^c} W^\theta(\mathbf{z}, i).$$

□

**Proof of Theorem B.1** Let  $I \in E_1$ . By Lemma A.2 for any  $i \in I^c$

$$P W(\mathbf{z}, i) \leq \rho_2 W(\mathbf{z}, i), |\mathbf{z}| \geq M.$$

Then using Jensen's inequality, we have

$$P W_\theta(\mathbf{z}) \leq \rho_2^\theta W_\theta(\mathbf{z}) \text{ if } |\mathbf{z}| \geq M. \quad (\text{B.10})$$

Define the constants

$$C_U := \sup \left\{ \frac{\prod_{i \in I} x_i^{\hat{p}_i}}{V(\mathbf{z})} : \mathbf{z} \in \mathbb{R}_+^{n, \circ} \times \mathbb{R}^{k_0} \right\} < \infty,$$

$$\varsigma := \frac{\delta_e^{\check{\theta}}}{C_U^\theta}$$

and the stopping time

$$\eta := \inf \{t \geq 0 : W_\theta(\mathbf{Z}(t)) \geq \varsigma\}.$$

Clearly, if  $W_\theta(\mathbf{z}) < \varsigma$ , then  $\eta > 0$  and for any  $i \in I^c$ , we get

$$X_i(t) \leq \delta_e, t \in [0, \eta]. \quad (\text{B.11})$$

Let

$$\tilde{W}_\theta(\mathbf{z}) := \varsigma \wedge W_\theta(\mathbf{z}).$$

We have from the concavity of  $x \mapsto x \wedge \varsigma$  that

$$\mathbb{E}_{\mathbf{Z}} \tilde{W}_{\theta}(\mathbf{Z}(T)) \leq \varsigma \wedge \mathbb{E}_{\mathbf{Z}} W_{\theta}(\mathbf{Z}(T)).$$

Let  $\tau$  be defined as in (A.20). By (B.10) we have that

$$\mathbb{E}_{\mathbf{Z}} \left[ \rho_2^{-\theta(\tau \wedge \eta \wedge n^* T_e)} W_{\theta}(\mathbf{Z}(\theta \gamma_b(\tau \wedge \xi \wedge n^* T_e))) \right] \leq W_{\theta}(\mathbf{z}), \mathbf{z} \in \mathcal{S}_+.$$

As a result for all  $\mathbf{z} \in \mathcal{S}_+$

$$\begin{aligned} W_{\theta}(\mathbf{z}) &\geq \mathbb{E}_{\mathbf{Z}} \left[ \rho_2^{-\theta(\tau \wedge \eta \wedge n^* T_e)} W_{\theta}(\mathbf{Z}(\tau \wedge \eta \wedge n^* T_e)) \right] \\ &\geq \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \tau\}} W_{\theta}(\mathbf{Z}(\tau)) \right] \\ &\quad + \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \eta\}} W_{\theta}(\mathbf{Z}(\eta)) \right] \\ &\quad + \rho_2^{-\theta(n^*-1)T_e} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T_e < \tau \wedge \eta \leq n^* T_e\}} W_{\theta}(\mathbf{Z}(\tau \wedge \eta)) \right] \\ &\quad + \rho_2^{-\theta n^* T_e} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \wedge \eta > n^* T_e\}} W_{\theta}(\mathbf{Z}(n^* T_e)) \right]. \end{aligned} \quad (\text{B.12})$$

By the strong Markov property of  $(\mathbf{Z}(t))$  and Proposition B.1 (which we can use because of (B.11)) we see that for all  $\mathbf{z} \in \mathcal{S}_+$

$$\begin{aligned} &\mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \tau\}} W_{\theta}(\mathbf{Z}(n^* T_e)) \right] \\ &\leq \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \tau\}} \exp \left( -\frac{1}{2} \theta r_e (n^* T_e - \tau) \right) W_{\theta}(\mathbf{Z}(\tau)) \right] \\ &\leq \exp \left( -\frac{1}{2} \theta r_e T_e \right) \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \tau\}} W_{\theta}(\mathbf{Z}(\tau)) \right]. \end{aligned} \quad (\text{B.13})$$

Similarly, by the strong Markov property of  $(\mathbf{Z}(t))$  and Lemma A.2, we obtain for any  $\mathbf{z} \in \mathcal{S}_+$  that

$$\begin{aligned} &\mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T_e < \tau \wedge \eta \leq n^* T_e\}} W_{\theta}(\mathbf{Z}(n^* T_e)) \right] \\ &\leq \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T_e < \tau \wedge \eta \leq n^* T_e\}} C_2^{\theta(n^* T_e - \tau \wedge \eta)} W_{\theta}(\mathbf{Z}(\tau \wedge \eta)) \right] \\ &\leq C_2^{\theta T_e} \mathbb{E}_{\mathbf{Z}} \left[ \mathbf{1}_{\{(n^*-1)T_e < \tau \wedge \eta \leq n^* T_e\}} W_{\theta}(\mathbf{Z}(\tau \wedge \eta)) \right]. \end{aligned} \quad (\text{B.14})$$

If  $W_\theta(\mathbf{z}) < \varsigma$  then applying (B.13), (B.14) and the inequality  $\tilde{W}_\theta(\mathbf{Z}(n^*T_e)) \leq W_\theta(\mathbf{Z}(n^*T_e \wedge \eta))$  to (B.12) yields

$$\begin{aligned} \tilde{W}_\theta(\mathbf{z}) &= W_\theta(\mathbf{z}) \geq \mathbb{E}_{\mathbf{z}} [\mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \tau\}} W_\theta(\mathbf{Z}(\tau))] \\ &\quad + \mathbb{E}_{\mathbf{z}} [\mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \eta\}} W_\theta(\mathbf{Z}(\eta))] \\ &\quad + \rho_2^{-\theta(n^*-1)T_e} \mathbb{E}_{\mathbf{z}} [\mathbf{1}_{\{(n^*-1)T_e < \tau \wedge \eta \leq n^*T\}} W_\theta(\mathbf{Z}(\tau \wedge \eta))] \\ &\quad + \rho_2^{-\theta n^*T_e} \mathbb{E}_{\mathbf{z}} [\mathbf{1}_{\{\tau \wedge \eta > n^*T_e\}} W_\theta(\mathbf{Z}(n^*T_e))] \\ &\geq \exp\left(\frac{1}{2}\theta r_e T_e\right) \mathbb{E}_{\mathbf{z}} [\mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \tau\}} W_\theta(\mathbf{Z}(n^*T_e))] \quad (\text{B.15}) \\ &\quad + \mathbb{E}_{\mathbf{z}} [\mathbf{1}_{\{\tau \wedge \eta \wedge (n^*-1)T_e = \eta\}} \tilde{W}_\theta(\mathbf{Z}(n^*T_e))] \\ &\quad + \rho_2^{-\theta(n^*-1)T_e} C_2^{-\theta T_e} \mathbb{E}_{\mathbf{z}} [\mathbf{1}_{\{(n^*-1)T_e < \tau \wedge \eta \leq n^*T_e\}} W_\theta(\mathbf{Z}(n^*T_e))] \\ &\quad + \rho_2^{-\theta n^*T_e} \mathbb{E}_{\mathbf{z}} [\mathbf{1}_{\{\tau \wedge \eta > n^*T_e\}} W_\theta(\mathbf{Z}(n^*T_e))] \\ &\geq \mathbb{E}_{\mathbf{z}} \tilde{W}_\theta(\mathbf{Z}(n^*T_e)) \quad (\text{since } \tilde{W}_\theta(\cdot) \leq W_\theta(\cdot)). \end{aligned}$$

Clearly, if  $W_\theta(\mathbf{z}) \geq \varsigma$  then

$$\mathbb{E}_{\mathbf{z}} \tilde{W}_\theta(\mathbf{Z}(n^*T_e)) \leq \varsigma = \tilde{W}_\theta(\mathbf{z}). \quad (\text{B.16})$$

As a result of (B.15), (B.16) and the Markov property of  $(\mathbf{Z}(t))$ , the sequence  $\{w_2(k) : k \geq 0\}$  where  $w_2(k) := \tilde{W}_\theta(\mathbf{Z}(kn^*T_e))$  is a supermartingale. Define the discrete stopping time

$$\eta^* := \inf\{k \in \mathbb{N} : W_\theta(\mathbf{Z}(kn^*T_e)) \geq \varsigma\}.$$

Moreover, we also deduce from (B.15) that

$$\mathbb{E}_{\mathbf{z}} \mathbf{1}_{\{\eta^* > 1\}} W_\theta(\mathbf{Z}(n^*T_e)) \leq \kappa_e W_\theta(\mathbf{z}), \mathbf{z} \in \mathcal{S}_+$$

where  $\kappa_e^{-1} = \min\left\{\rho_2^{-\theta(n^*-1)T_e} C_2^{-\theta T_e}, \exp\left(\frac{1}{2}\theta r_e T_e\right), \rho_2^{-\theta n^*T_e}\right\} > 1$ . As a result,  $\{w_3(k) : k \geq 0\}$  with

$$w_3(k) := \kappa_e^{-k} \mathbf{1}_{\{\eta^* > k\}} W_\theta(\mathbf{Z}(kn^*T_e))$$

is also a supermartingale. For any  $\varepsilon \in (0, 1)$ , if  $W_\theta(\mathbf{z}) \leq \varsigma \varepsilon$  we have

$$\mathbb{E}_{\mathbf{z}} w_2(k \wedge \eta^*) \leq \mathbb{E}_{\mathbf{z}} w_2(0) = W_\theta(\mathbf{z}) \leq \varsigma \varepsilon, k \geq 0. \quad (\text{B.17})$$

Subsequently, (B.17) combined with the Markov inequality and the fact that  $w_2(\eta) = W_\theta(\mathbf{Z}(\eta n^*T_e)) \geq \varsigma$  yields

$$\mathbb{P}_{\mathbf{z}}\{\eta^* < k\} \leq \varsigma^{-1} \mathbb{E}_{\mathbf{z}} w_2(k \wedge \eta^*) \leq \varepsilon, \text{ if } k \in \mathbb{N}, W_\theta(\mathbf{z}) \leq \varsigma \varepsilon.$$

Next, let  $k \rightarrow \infty$  to get

$$\mathbb{P}_{\mathbf{z}}\{\eta^* < \infty\} \leq \varepsilon \text{ if } W_{\theta}(\mathbf{z}) \leq \varsigma \varepsilon. \quad (\text{B.18})$$

Let  $\kappa_3 \in (\kappa_e, 1)$ . Since  $P_t \tilde{W}_{\theta}(\mathbf{z}) \leq C_2^{\theta t} \tilde{W}_{\theta}(\mathbf{z})$  for any  $t > 0$ ,  $\mathbf{z} \in \mathcal{S}_+$ , we have that

$$\mathbb{P}_{\mathbf{z}}\{\tilde{W}_{\theta}(\mathbf{Z}(t)) \leq c, t = 1, \dots, n^* T_e\} \leq \frac{n^* T_e C_2^{\theta n^* T_e} \tilde{W}_{\theta}(\mathbf{z})}{c}.$$

The last inequality, the Markov property of  $\mathbf{Z}(t)$ , and the fact that  $(w_3(k))_{k \geq 0}$  is a supermartingale imply that for any  $\mathbf{z} \in \mathcal{S}_+$

$$\begin{aligned} \mathbb{P}_{\mathbf{z}}\left\{\mathbf{1}_{\{\eta^* > k\}} \tilde{W}_{\theta}(\mathbf{Z}(t)) \leq \left(\frac{\kappa_e}{\kappa_3}\right)^{-k}, t \in [kn^* T_e + 1, (k+1)n^* T_e]\right\} \\ \leq n^* T_e C_2^{\theta n^* T_e} \mathbb{E}_{\mathbf{z}} w_3(k) \frac{\kappa_e^{-k}}{\kappa_3^{-k}} \leq n^* T_e C_2^{\theta n^* T_e} \frac{\mathbb{E}_{\mathbf{z}} w_3(0)}{\kappa_3^{-k}} \leq n^* T_e C_2^{\theta n^* T_e} \frac{W_{\theta}(\mathbf{z})}{\kappa_3^{-k}} \end{aligned}$$

Since  $\sum_k \left(n^* T_e C_2^{\theta n^* T_e} \frac{W_{\theta}(\mathbf{z})}{\kappa_3^{-k}}\right) < \infty$ , we deduce from the Borel–Cantelli lemma that

$$\mathbb{P}_{\mathbf{z}}\left\{\limsup_{k \rightarrow \infty} \sup_{t \in [kn^* T_e + 1, (k+1)n^* T_e]} \mathbf{1}_{\{\eta^* > k\}} \tilde{W}_{\theta}(\mathbf{Z}(t)) \left(\frac{\kappa_3}{\kappa_e}\right)^k < 1\right\} = 1.$$

This and (B.18) imply that if  $W_{\theta}(\mathbf{z}) \leq \varsigma \varepsilon$  then

$$\mathbb{P}_{\mathbf{z}}\left\{\limsup_{t \rightarrow \infty} \left(\left(\frac{\kappa_3}{\kappa_e}\right)^{\frac{t}{n^* T_e + 1}} W_{\theta}(\mathbf{Z}(t))\right) < 1\right\} \geq 1 - \varepsilon.$$

Since  $\sup_{\mathbf{z} \in \mathbb{R}^n \times \mathbb{R}^{\kappa_0}} \frac{x_i^{\theta \tilde{p}}}{W_{\theta}(\mathbf{z})} < \infty$ ,  $i \in I^c$ , we can easily obtain the extinction result from Theorem 2.5.  $\square$

## Appendix C. Robustness proofs

Let  $\circ$  denote the element-wise product and  $\mathbf{1}_n$  be the vector in  $\mathbb{R}^n$  whose components are all 1. Assume that the function  $V$  from Assumption A3) satisfies the robust estimate

$$\mathbb{E}\left[V(\mathbf{x}^{\top}(F(\mathbf{z}, \xi) \circ (\mathbf{1}_n + \tilde{\varepsilon}_1)), G(\mathbf{z}, \xi) + \tilde{\varepsilon}_2))h(\mathbf{z}, \xi)\right] \leq \rho V(\mathbf{z}) + C$$

for any vectors  $\tilde{\varepsilon}_1, \tilde{\varepsilon}_2 \in \mathbb{R}^n$  such that  $|\tilde{\varepsilon}_1| \vee |\tilde{\varepsilon}_2| < \delta$ . Note that

$$\tilde{h}(\mathbf{z}, \xi) = \max_{i=1}^n \left\{ \tilde{F}_i(\mathbf{z}, \xi), \frac{1}{\tilde{F}_i(\mathbf{z}, \xi)} \right\}^{\gamma_3}.$$

Consequently  $\tilde{h}(\mathbf{z}, \xi) \leq e^\delta h(\mathbf{z}, \xi)$  and

$$\mathbb{E} \left[ V(\mathbf{x}^\top (\tilde{F}(\mathbf{z}, \xi)), \tilde{G}(\mathbf{z}, \xi)) \tilde{h}(\mathbf{z}, \xi) \right] \leq \rho V(\mathbf{z}) + C$$

if  $\delta > 0$  is sufficiently small. This shows that there exist  $C_2 > 0$ ,  $\rho_2 \in (0, 1)$ ,  $\gamma_4 \in (0, \gamma_3)$  such that

$$\mathbb{E}_{\mathbf{z}} V \left( \tilde{\mathbf{Z}}(1) \prod_{i=1}^n \tilde{X}_i^{p_i}(1) \right) \leq (\mathbf{1}_{\{|\mathbf{z}| < M\}} (C_2 - \rho_2) + \rho_2) V^{\frac{1}{2}}(\mathbf{z}) \prod_{i=1}^n x_i^{p_i}$$

for any  $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{R}^n$  satisfying

$$|\mathbf{p}|_1 := \sum |p_i| \leq \gamma_4.$$

Analogously to Lemma A.1 one can show the following uniform bound

$$\sup_{|\mathbf{z}| \leq M, t \in \mathbb{N}} \left( \mathbb{E}_{\mathbf{z}} V(\tilde{\mathbf{Z}}(t)) + \mathbb{E}_{\mathbf{z}} \tilde{h}(\tilde{\mathbf{Z}}(t), \xi(t)) \right) = K_M < \infty \quad (\text{C.1})$$

when  $\delta$  is sufficiently small.

Slight changing in the factor on the right hand side of A.5, we have

**Lemma C.1** *Suppose that Assumption A.1 holds. Let  $\mathbf{p}$  and  $r^*$  be as in (A.4). There exists a  $\tilde{T}^* > 0$  such that, for any  $T > \tilde{T}^*$ ,  $\mathbf{z} \in \mathcal{S}_0$ ,  $\|\mathbf{z}\| \leq M$  one has*

$$\sum_{t=0}^T \mathbb{E}_{\mathbf{z}} \left( \ln V(\mathbf{Z}(t+1)) - \ln V(\mathbf{Z}(t)) - \sum p_i \ln F_i(\mathbf{Z}(t), \xi(t)) \right) \leq -1.5r^*(T+1). \quad (\text{C.2})$$

On the other hand, it obviously follows from (2.11) that for any  $\varepsilon > 0$ ,  $M > 0$ ,  $T > 0$ ,  $n_0 > 0$ , we have

$$\mathbb{E}_{\mathbf{z}} \left\{ \mathbf{1}_{\{|\tilde{\mathbf{Z}}(t)| \vee |\mathbf{Z}(t)| < n_0\}} |\tilde{\mathbf{Z}}(t) - \mathbf{Z}(t)| \right\} \leq \varepsilon, \mathbf{z} \leq M, t \in \{0, \dots, T\}$$

when  $\delta$  is sufficiently small. This and the uniform integrability (C.1) imply that, for any  $\varepsilon$ ,  $T > 0$ , there exists  $\delta > 0$  such that for any  $\delta$ -perturbation of (3.1), we have

$$\begin{aligned} \mathbb{E}_{\mathbf{z}} |\ln \tilde{F}_i(\tilde{\mathbf{Z}}(t)) - \ln F_i(\mathbf{Z}(t))| &\leq \mathbb{E}_{\mathbf{z}} |\ln \tilde{F}_i(\tilde{\mathbf{Z}}(t)) - \ln \tilde{F}_i(\mathbf{Z}(t))| \\ &+ \mathbb{E}_{\mathbf{z}} |\ln \tilde{F}_i(\mathbf{Z}(t)) - \ln F_i(\mathbf{Z}(t))| \varepsilon, \mathbf{z} \in \mathcal{S}_0, |\mathbf{z}| \leq M, t \in \{0, \dots, T\}. \end{aligned} \quad (\text{C.3})$$

By virtue of (C.2) and (C.3), for sufficiently small  $\delta$ ,

$$\sum_{t=0}^T \mathbb{E}_{\mathbf{z}} \left( \ln V(\tilde{\mathbf{Z}}(t+1)) - \ln V(\tilde{\mathbf{Z}}(t)) - \sum p_i \ln F_i(\tilde{\mathbf{Z}}(t), \xi(t)) \right) \leq -r^*(T+1) \quad (\text{C.4})$$

for any  $\mathbf{z} \in S_0$ ,  $|\mathbf{z}| \leq M$  and  $T \geq \tilde{T}^*$ . With this estimate, it follows from the arguments in Appendix A that Theorem A.1 holds for  $\tilde{\mathbf{Z}}(t)$ . Similarly, Theorem 2.5 and Theorem 2.6 hold.

## Appendix D. Proof of Lemma 4.1

**Lemma D.1** Suppose  $\{(E_i(t), S(t))_{i=1, \dots, n}\}$ ,  $t \in \mathbb{N}$  is a sequence of  $n+1$ -dimensional random variables, i.i.d. over  $t$  such that  $\mathbb{E}[S(t)e^{E_j(t)}]^2 < \infty$ . Then the model given by (2.4) and (4.11) satisfies Assumption 2.1 by taking a small enough  $\gamma_3 > 0$ , and

$$V(\mathbf{z}) = \sum_j z_j + 1.$$

Assumption 2.2 holds with

$$\phi(\mathbf{z}) = \delta V(\mathbf{z})$$

for some  $\delta > 0$ . Moreover, if the support of  $\ln S(t) + \max_j (E_j(t) \ln \delta_j)$  contains values less than 0 then the boundary is accessible.

**Proof** We denote by  $K$  below a positive generic constant. Using (4.11)

$$\sum_j N_j(t) e^{E_j(t) - D_j(t)} \leq \sum_j N_j(t) e^{E_j(t)} \exp\{-\alpha_j N_j(t) e^{E_j(t)}\} < K < \infty. \quad (\text{D.1})$$

On the other hand

$$\left(1 - \delta_j + S(t) e^{E_j(t) - D_j(t)}\right) \vee \left(1 - \delta_j + S(t) e^{E_j(t) - D_j(t)}\right)^{-1} \leq K + \sum_j S(t) e^{E_j(t) - D_j(t)}$$

Using this and the inequality

$$(a+b)^\gamma \leq 2^\gamma (a^\gamma + b^\gamma)$$

we see that for any  $\varepsilon > 0$  there is  $\gamma_3 > 0$  sufficiently small such that

$$\begin{aligned} h(t) &:= \max_j \left( \left(1 - \delta_j + S(t) e^{E_j(t) - D_j(t)}\right)^{\gamma_3} \vee \left(1 - \delta_j + S(t) e^{E_j(t) - D_j(t)}\right)^{-\gamma_3} \right) \\ &\leq \left(1 + \varepsilon + \left(S(t) \sum e^{E_j(t) - D_j(t)}\right)^{\gamma_3}\right) \end{aligned}$$



Using (D.1) and the fact that  $\mathbb{E}S(t) < \infty$  we get

$$\mathbb{E}^t \sum_j \left( (1 - \delta_j) N_j(t) (1 + \varepsilon + (S(t) e^{E_j(t) - D_j(t)})^{\gamma_3}) \right) \leq (1 - 0.5\check{d}) \sum_j N_j(t) + K$$

where  $\mathbb{E}^t[\cdot] := \mathbb{E}[\cdot \mid N_1(t), \dots, N_n(t)]$  and  $\check{d} = \min\{\delta_j\}$ . Since  $N_j(t) e^{E_j(t)} \exp\{-a_j(1 + \gamma_3)N_j(t) e^{E_j(t)}\}$  is bounded above by a nonrandom constant, then

$$N_j(t) e^{(1+\gamma_3)(E_j(t) - D_j(t))} \leq e^{\gamma_3 E_j(t)} N_j(t) e^{E_j(t)} \exp\{-a_j(1 + \gamma_3)N_j(t) e^{E_j(t)}\} < K e^{\gamma_3 E_j(t)}$$

Therefore, since  $\mathbb{E}[S^2(t) e^{\gamma_3 E_j(t)}] < \infty$  we have

$$\mathbb{E}^t S^{1+\gamma_3}(t) N_j(t) e^{(1+\gamma_3)(E_j(t) - D_j(t))} \leq K \mathbb{E}^t \left( S^{1+\gamma_3}(t) e^{\gamma_3 E_j(t)} \right) \leq K_2.$$

As a result,

$$\begin{aligned} \mathbb{E}^t \sum_j \left( 1 - \delta_j + S(t) e^{E_j(t) - D_j(t)} \right) N_j(t) h(t) &\leq \sum_j (1 - 0.5\delta_j) N_j(t) \\ &+ K_2 \leq (1 - 0.25\check{d}) \sum_j N_j(t) + K_2. \end{aligned}$$

This implies that Assumption A3) is satisfied with  $V(\mathbf{z}) = \sum_j N_j + 1 = \sum_j z_j + 1$  and small  $\gamma_3 > 0$ . Note that

$$\sum_j N_j(t+1) - P_1 \sum_j N_j(t) = S(t) \sum_j e^{E_j(t) - D_j(t)} N_j(t) - \mathbb{E}^t S(t) \sum_j e^{E_j(t) - D_j(t)} N_j(t)$$

This shows that

$$\begin{aligned} \mathbb{E}^t \left( \sum_j N_j(t+1) - P_1 \sum_j N_j(t) \right)^2 &\leq \mathbb{E}^t \left( S(t) \sum_j e^{E_j(t) - D_j(t)} N_j(t) \right)^2 \\ &\leq K_3^2 \mathbb{E}^t \left[ S(t) e^{E_j(t)} \right]^2 = K_3^2 \mathbb{E}^t \left[ S(t) e^{E_j(t)} \right]^2 \end{aligned}$$

since  $\sum_j e^{-D_j(t)} N_j(t)$  is bounded by a constant  $K_3$ . If  $\mathbb{E}[S(t) e^{E_j(t)}]^2 < \infty$  then Assumption 2.2 is satisfied with  $\phi(\mathbf{z}) = \delta(|\mathbf{z}| + 1)$  for some  $\delta > 0$ .

Moreover, if support of  $\ln S(t) + \max_j \{E_j(t) \ln \delta_j\}$  contains values less than 0 then it is clear that the boundary is accessible since when  $\ln S(t) + \max_j \{E_j(t) \ln \delta_j\}$  is less than a negative constant,  $N_j(t+1) \leq \rho N_j(t)$  for  $\rho < 1$ .  $\square$

## References

- Abrams P (1984) Variability in resource consumption rates and the coexistence of competing species. *Theor Popul Biol* 25:106–124
- Abrams P, Holt R (2002) The impact of consumer-resource cycles on the coexistence of competing consumers. *Theor Popul Biol* 62:281–295
- Andrewartha H, Birch L (1954) The distribution and abundance of animals. University of Chicago Press, Chicago
- Armstrong RA, McGehee R (1980) Competitive exclusion. *Am Nat* 115(2):151–170
- Bartlett MS (1960) Stochastic population models in ecology and epidemiology. Methuen, London
- Benaïm M (2018) Stochastic persistence. arXiv preprint [arXiv:1806.08450](https://arxiv.org/abs/1806.08450)
- Benaïm M, Hofbauer J, Sandholm WH (2008) Robust permanence and impermanence for stochastic replicator dynamics. *J Biol Dyn* 2(2):180–195
- Benaïm M, Schreiber SJ (2019) Persistence and extinction for stochastic ecological models with internal and external variables. *J Math Biol* 79:393–431
- Cantrell RS, Cosner C, Lou Y (2012) Evolutionary stability of ideal free dispersal strategies in patchy environments. *J Math Biol* 65(5):943–965
- Chesson P (1978) Predator–prey theory and variability. *Annu Rev Ecol Syst* 9(1):323–347
- Chesson P (1994) Multispecies competition in variable environments. *Theor Popul Biol* 45(3):227–276
- Chesson P (2000a) General theory of competitive coexistence in spatially-varying environments. *Theor Popul Biol* 58(3):211–237
- Chesson P (2000b) Mechanisms of maintenance of species diversity. *Annu Rev Ecol Syst* 31:343–66
- Chesson P (2003) Quantifying and testing coexistence mechanisms arising from recruitment fluctuations. *Theor Popul Biol* 64(3):345–357
- Chesson P (2008) Quantifying and testing species coexistence mechanisms. *Fundacion BBVA, Bilbao*, pp 119–164
- Chesson P (2018) Updates on mechanisms of maintenance of species diversity. *J Ecol* 106(5):1773–1794
- Chesson P (2019) Chesson’s coexistence theory: comment. *Ecology*. <https://doi.org/10.1002/ecy.2851>
- Chesson P (2020) Species coexistence, book section 2. Oxford University Press, Oxford, pp 5–27
- Chesson P, Huntly N (1997) The roles of harsh and fluctuating conditions in the dynamics of ecological communities. *Am Nat* 150(5):519–553
- Chesson P, Huntly NJ, Roxburgh SH, Pantastico-Caldas M, Facelli JM (2013) The storage effect: definition and tests in two plant communities, book section 2. Cambridge University Press, New York, pp 11–40
- Chesson P, Kuang JJ (2010) The storage effect due to frequency-dependent predation in multispecies plant communities. *Theor Popul Biol* 78(2):148–164
- Chesson PL (1982) The stabilizing effect of a random environment. *J Math Biol* 15(1):1–36
- Chesson PL (1983) Coexistence of competitors in a stochastic environment: the storage effect. *Lect Notes Biomath* 52:188–198
- Chesson PL (1985) Coexistence of competitors in spatially and temporally varying environments: a look at the combined effects of different sorts of variability. *Theor Popul Biol* 28:263–287
- Chesson PL (1989) A general model of the role of environmental variability in communities of competing species. *Am Math Soc Lect Math Life Sci* 20:97–123
- Chesson PL, Ellner S (1989) Invasibility and stochastic boundedness in monotonic competition models. *J Math Biol* 27(2):117–138
- Chesson PL, Huntly N (1988) Community consequences of life-history traits in a variable environment. *Ann Zool Fenn* 25:5–16
- Chesson PL, Warner RR (1981) Environmental variability promotes coexistence in lottery competitive systems. *Am Nat* 117(6):923–943
- Connell J (1978) Diversity in tropical rain forests and coral reefs. *Science* 199:1302–1310
- Diaconis P, Freedman D (1999) Iterated random functions. *SIAM review* 41(1):45–76
- Durrett R, Remenik D (2012) Evolution of dispersal distance. *J Math Biol* 64(4):657–666
- Ellner S (1989) Convergence to stationary distributions in two-species stochastic competition models. *J Math Biol* 27(4):451–462
- Evans SN, Hening A, Schreiber SJ (2015) Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments. *J Math Biol* 71(2):325–359
- Evans SN, Ralph PL, Schreiber SJ, Sen A (2013) Stochastic population growth in spatially heterogeneous environments. *J Math Biol* 66(3):423–476

- Fudenberg D, Harris C (1992) Evolutionary dynamics with aggregate shocks. *J Econ Theory* 57(2):420–441
- Garay BM, Hofbauer J (2003) Robust permanence for ecological differential equations, minimax, and discretizations. *SIAM J Math Anal* 34(5):1007–1039
- Gonzalez A, Holt RD (2002) The inflationary effects of environmental fluctuations in source–sink systems. *Proc Natl Acad Sci* 99(23):14872–14877
- Gyllenberg M, Jiang J, Niu L, Yan P (2019) On the dynamics of multi-species ricker models admitting a carrying simplex. *J Differ Equ Appl* 25:1489–1530
- Hardin DP, Takáč P, Webb GF (1988a) Asymptotic properties of a continuous-space discrete-time population model in a random environment. *J Math Biol* 26(4):361–374
- Hardin DP, Takáč P, Webb GF (1988b) A comparison of dispersal strategies for survival of spatially heterogeneous populations. *SIAM J Appl Math* 48(6):1396–1423
- Hardin DP, Takáč P, Webb GF (1990) Dispersion population models discrete in time and continuous in space. *J Math Biol* 28(1):1–20
- Hastings A (1983) Can spatial variation alone lead to selection for dispersal? *Theor Popul Biol* 24(3):244–251
- Hening A (2021) Coexistence, extinction, and optimal harvesting in discrete-time stochastic population models. *J Nonlinear Sci* 31:1–50
- Hening A, Nguyen DH (2018a) Coexistence and extinction for stochastic Kolmogorov systems. *Ann Appl Probab* 28(3):1893–1942
- Hening A, Nguyen DH (2018b) Persistence in stochastic Lotka–Volterra food chains with intraspecific competition. *Bull Math Biol* 80(10):2527–2560
- Hening A, Nguyen DH (2018c) Stochastic Lotka–Volterra food chains. *J Math Biol* 77(1):135–163
- Hening A, Nguyen DH, Yin G (2018) Stochastic population growth in spatially heterogeneous environments: the density-dependent case. *J Math Biol* 76(3):697–754
- Hening A, Nguyen D, Schreiber SJ (2021) A classification of the dynamics of three-dimensional stochastic ecological systems. *Ann Appl Probab*
- Hofbauer J (1981) A general cooperation theorem for hypercycles. *Monatshefte für Mathematik* 91(3):233–240
- 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
- Hofbauer J, Imhof LA (2009) Time averages, recurrence and transience in the stochastic replicator dynamics. *Ann Appl Probab* 19(4):1347–1368
- Hutchinson GE (1961) The paradox of the plankton. *Am Nat* 95:137–145
- Imhof LA (2005) The long-run behavior of the stochastic replicator dynamics. *Ann Appl Probab* 15(1B):1019–1045
- Kendall D (1948) On the generalized birth and death process. *Ann Math Stat* 19:1–15
- Kuang JJ, Chesson P (2008) Predation-competition interactions for seasonally recruiting species. *Am Nat* 171(3):E119–E133
- Kuang JJ, Chesson P (2009) Coexistence of annual plants: generalist seed predation weakens the storage effect. *Ecology* 90(1):170–182
- Kuang JJ, Chesson P (2010) Interacting coexistence mechanisms in annual plant communities: frequency-dependent predation and the storage effect. *Theor Popul Biol* 77(1):56–70
- Kurtz TG (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes. *J Appl Probab* 7(1):49–58
- Li L, Chesson P (2016) The effects of dynamical rates on species coexistence in a variable environment: the paradox of the plankton revisited. *Am Nat* 188(2):E46–E58
- May RM (1974) *Stability and complexity in model ecosystems*, 2nd edn. Princeton University Press, Princeton, NJ
- Meyn SP, Tweedie RL (1992) Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv Appl Probab* 24(3):542–574
- Ripa J, Ives AR (2003) Food web dynamics in correlated and autocorrelated environments. *Theor Popul Biol* 64(3):369–384
- Roth G, Schreiber SJ (2014) Persistence in fluctuating environments for interacting structured populations. *J Math Biol* 69(5):1267–1317
- Roy M, Holt RD, Barfield M (2005) Temporal autocorrelation can enhance the persistence and abundance of metapopulations comprised of coupled sinks. *Am Nat* 166(2):246–261
- Ryals B, Sacker RJ (2015) Global stability in the 2d Ricker equation. *J Differ Equ Appl* 21(11):1068–1081

- Ryals B, Sacker RJ (2016) Global stability in the 2d ricker equation revisited. *Discrete Contin Dyn Syst B* 22(2):585
- Schmidt KA (2004) Site fidelity in temporally correlated environments enhances population persistence. *Ecol Lett* 7(3):176–184
- Schreiber SJ (2010) Interactive effects of temporal correlations, spatial heterogeneity and dispersal on population persistence. *Proc R Soc Lond B Biol Sci* 277:1907–1914
- Schreiber SJ (2012) Persistence for stochastic difference equations: a mini-review. *J Differ Equ Appl* 18(8):1381–1403
- Schreiber SJ, Benaïm M, Atchadé KAS (2011) Persistence in fluctuating environments. *J Math Biol* 62(5):655–683
- Strong DR (1986) Density vagueness: abiding the variance in the demography of real populations. Harper & Row Publishing, New York
- Turelli M (1978) Does environmental variability limit niche overlap? *Proc Natl Acad Sci USA* 75:5085–5089
- Turelli M (1981) Niche overlap and invasion of competitors in random environments i: models without demographic stochasticity. *Theor Popul Biol* 20:1–56
- Turelli M, Gillespie JH (1980) Conditions for the existence of stationary densities for some two-dimensional diffusion processes with applications in population biology. *Theor Popul Biol* 17(2):167–189
- Yuan C, Chesson P (2015) The relative importance of relative nonlinearity and the storage effect in the lottery model. *Theor Popul Biol* 105:39–52

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