Stochastic Persistence*
(Part I)

Michel Benaïm†

June 28, 2018

Abstract

Let $(X_t)_{t \geq 0}$ be a continuous time Markov process on some metric space M, leaving invariant a closed subset $M_0 \subset M$, called the *extinction set*. We give general conditions ensuring either

Stochastic persistence (Part I) : Limit points of the occupation measure are invariant probabilities over $M_+ = M \setminus M_0$; or

Extinction (Part II) : $X_t \to M_0$ a.s.

In the persistence case we also discuss conditions ensuring the a.s convergence (respectively exponential convergence in total variation) of the occupation measure (respectively the distribution) of (X_t) toward a unique probability on M_+.

These results extend and generalize previous results obtained for various stochastic models in population dynamics, given by stochastic differential equations, random differential equations, or pure jump processes.

Keywords Stochastic persistence, Lyapunov and average Lyapunov functions, Markov processes, Ergodicity

Contents

*Revised version of the unpublished preprint "Stochastic Persistence, April 2014"

†Institut de Mathématiques, Université de Neuchâtel, Rue Emile-Argand, Neuchâtel, Suisse-2000. (michel.benaim@unine.ch).

1 Introduction

An important issue in mathematical ecology and population biology is to find out under which conditions a collection of interacting species can coexist over long periods of time. A similar question, in mathematical models of disease dynamics, is to understand whether or not a disease will be endemic (i.e persist in the population) or go extinct. The mathematical investigation of these types of questions began with the early work of Freedman and Waltman [29], Gard [32, 33], Gard and Hallam [34], Schuster Sigmund and Wolff [62], among others, in the late 1970s, laying the foundation of what is now called the (deterministic) *mathematical theory of persistence.* The theory developed rapidly the past 35 years using the available tools from dynamical system theory. The recent books by Smith and Thieme [73]; Zhao and Borwein [78] provide a comprehensive introduction to the theory as well as numerous examples and references.

For (most of) deterministic models, *persistence* amounts to say that there exists an attractor bounded away from the *extinction states* (i.e the subset of the states space where the abundance of one or group of the species vanishes). When this attractor is global, meaning that its basin of attraction includes all non-extinction states, the system is called *uniformly persistent* or *permanent* [62, 41].

Beside biotic interactions, environmental fluctuations play a key role in population dynamics. In order to take into account these fluctuations and to understand how they may affect persistence, one approach is the study of uniform persistence for non-autonomous difference or differential equations [75, 60, 73]. Another is to consider *systems subjected to environmental random perturbations.* Classical examples include *ecological stochastic differential equations* (see e.g the classical paper by Turreli [76] or [49]).

$$dx_i = x_i[F_i(x)dt + \sum_{j=1}^{m} \Sigma_i^j(x)dB_t^j],\ i = 1 \dots n \tag{1}$$

where $(B_t^1, \dots, B_t^m)$ is a standard m-dimensional Brownian motion; and *ecological stochastic equations driven by a Markov chain*

$$\frac{dx_i}{dt} = x_i(t)F_i^{J(t)}(x(t)),\ i = 1 \dots n \tag{2}$$

where $J(t) \in \{1, \dots, m\}$ is a continuous time Markov chain - or more generally, a continuous time Markov chain controlled by $(x(t))$ - taking values in a finite set representing different possible environments. Both (1) and (2) are Markov processes defined on $M = \mathbb{R}_+^n$ (respectively $\mathbb{R}_+^n \times \{1, \ldots, m\}$) and describe the evolution of n interacting species characterized by their abundances $x_1, \ldots, x_n$. The extinction set is the boundary $M_0 = \partial \mathbb{R}_+^n$ (respectively $\mathbb{R}_+^n \times \{1, \ldots, m\}$.

Generalizing upon these models we will consider here a continuous time Markov process (X_t) living in some metric space M and leaving invariant a closed subset $M_0 \subset M$, called the *extinction set*. That is

$$X_0 \in M_0 \Leftrightarrow X_t \in M_0 \text{ for all } t \geq 0.$$

Observe that, when $X_0 \in M^+ := M \setminus M_0$, (X_t) is never absorbed by M_0 and extinction can only occur *asymptotically*. The long term behavior of the process is then completely different from the behavior of a process that would be absorbed (or killed) in *finite time* (see e.g the beautiful survey by Villemonais and Méléard [61] for a discussion of such processes). While extinction occurs in finite time for most "realistic" finite population models, this extinction may be proceeded by long-term term transients when habitat sizes are sufficiently large. Hence, under this assumption, one can ignore the effects of *demographic stochasticity* (i.e. finite population effects) and focus on models with only *environmental stochasticity* where extinction can only be asymptotic. The recent survey paper by Schreiber [71] discusses these distinctions. Since the early observation by Hutchinson [45] that temporal fluctuations of the environment can favor coexistence of species despite very limited resources, the effect of environmental stochasticity has been widely explored in the ecology literature, especially through the influence of Chesson and his coauthors [19, 18, 14, 17].

For deterministic models given by ecological differential equations - that is equation (1) with $\Sigma_i^j = 0$ or (2) with $m = 1$ - general sufficient conditions ensuring permanence or extinction (and generalizing many of the existing results), were derived by Hofbauer, Schreiber, and their co-authors in a series of papers [68, 31, 44]. They rely on the existence of a suitable *average Lyapunov function*, a powerful notion introduced by Hofbauer [41] in the early 1980s.

The central idea of the present paper is to define a similar object for Markov processes. First attempts in this directions include [6] dealing with small random perturbations of deterministic systems (i.e (1) with small Σ_i^j) and later [70] for more general systems on compact state spaces (see also [11] and [67] for discrete time models). The results in [6, 70] have been recently generalized by Hening and Nguyen [37] allowing to treat (1) in full generality provided the diffusion term is non-degenerate.

In rough terms, our key assumption will be that there exist real valued continuous functions V and LV defined on M^+ with $V \geq 0$ (and typically $V(x) \to \infty$ as $x \to M_0$) such that

(a) The process

$$M_t = V(X_t) - V(X_0) - \int_0^t LV(X_s)ds, t \geq 0$$

is a martingale for all $X_0 \in M^+$;

(b) LV extends continuously to a function H defined on all M.

In the deterministic case where X_t is solution to an ordinary differential equation, say $\dot{X} = G(X)$, then $LV = \langle G, \nabla V \rangle$, $M_t = 0$, and we recover Hofbauer's notion of average Lyapunov function.

Associated to (V, H) are the *H-exponents*

$$\Lambda^-(H) = -\sup \int H(x)\mu(dx), \Lambda^+(H) = -\inf \int H(x)\mu(dx)$$

where the supremum (respectively infimum) is taken over the set of ergodic measures for (X_t) supported by M_0. The sign of these exponents determine the behavior of the process near the extinction set. We will show that (under certain technical assumptions):

- (Part I). If $\Lambda^-(H)$ is positive, then
  - The process is *stochastically persistent*, meaning that every limit point Π of its empirical measure

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

    is almost surely an invariant measure on M^+. That is $\Pi(M^+) = 1$.
  - Under further irreducibility condition, such an invariant measure is unique and the law of (X_t) converges to Π possibly at an exponential rate.

- (Part II). If $\Lambda^+(H)$ is negative, then $X_t \to M_0$ at rate

$$\liminf_{t\to\infty} \frac{V(X_t)}{t} \geq -\Lambda^+(H)$$

This paper is a fully revised and extended version of the unpublished notes [4], accompanying the Bernoulli lecture given by the author at the Centre Interfacultaire Bernoulli in october 2014. Some of the ideas contained in these notes, have been already used in a few papers ([10, 37, 12, 38] devoted to the analysis of certain ecological models. The present version has greatly benefitted from these papers. In particular, the beautiful analysis of the ecological sde (1) conducted by Hening and Nguyen [37] has helped to formulate conditions to deal with the situation where the extinction set is noncompact. Joint work with Edouard Strickler [12] has helped to understand how the general results here can be applied to the situation where the extinction set is no longer the boundary of the state space but an equilibrium point (a situation which naturally occurs in epidemic model), which after a natural change of variables, becomes a sphere. Discussions with Joseph Hofbauer and Sebastian Schreiber over the recent years have been particularly influential.

Outline The organization of Part I is as follows. Section 2 introduces the notation and the main assumptions, ensuring in particular tightness of empirical measures. Section 3 describes some motivating examples. Section 4 contains the main results: the persistence theorem (Theorem 4.4), conditions ensuring uniqueness of a persistent measure, convergence to this measure (Proposition 4.7 and Theorem 4.9), and under additional assumptions, exponential convergence (Theorems 4.10 and 4.12). Section 5 applies these results to ecological SDEs (equation (1)) including degenerate ones. As an illustration, Section 5.2 analyzes a Rosenzweig-MacArthur model where the prey variable (but not the predator variable) is subjected to some small Brownian perturbation. Section 6 considers random ODEs driven by a Markov Chain (equation (2)) and, as an illustration, fully analyzes in Section 6.1 a 3-dimensional process obtained by random switching between two May and Leonard vector fields. This provides an example for which the extinction set is not simply the boundary of the state space, but here the union of this boundary and an invariant line. The stochastic persistence results combined with known results on competitive systems (in particular the theory of carrying simplices) allow to give precise conditions ensuring the existence of a unique persistent measure, absolutely continuous with respect to Lebesgue, and to characterize its topological support as the cell bordered by the carrying simplices of the two vector fields. Section 7 contains the proof of the persistence Theorem and Section 8 the proof of the exponential convergence results. Section 9.2 is an appendix gathering some folklore results and their proofs.

2 Notation and hypotheses

Let (M,d) be a locally compact Polish space (e.g $\mathbb{R}^n, \mathbb{R}^n_+$ with the usual distance metric), equipped with its Borel σ-algebra $\mathcal{B}(M)$. We denote by $(\mathcal{M}_b(M), ||\cdot||)$ the Banach space of all real-valued bounded measurable functions on M under the *sup-norm* metric $||\cdot||$ and $C_b(M)$ (respectively $C_0(M)$) the Banach (sub)space of real-valued bounded continuous functions on M (respectively real valued continuous functions vanishing at infinity). For any set $A \subset M$ we let $\mathbf{1}_A$ denote the indicator function of A. A generic non-negative constant is noted cst. We let $\mathcal{P}(M)$ denote the space of probability measures on $\mathcal{B}(M)$ equipped with the the topology of weak convergence. For $\mu \in \mathcal{P}(M)$ and $f \in \mathcal{M}_b(M)$, we write $\mu f = \int_M f(x)\mu(dx)$. Recall that a sequence $(\mu_n)_{n\geq 1} \subset \mathcal{P}(M)$ is said to converge weakly to $\mu \in \mathcal{P}(M)$, written $\mu_n \Rightarrow \mu$, if for all $f \in C_b(M),\ \mu_n f \to \mu f$.

Throughout the paper, we assume given a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, a complete right continuous filtration $(\mathcal{F}_t)$, and a family of *cad-lag* Markov processes $\{(X_t^x)_{t\geq 0},\, x \in M\}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$. By this we mean that

(i) For all $x \in M$ X_t^x is a $M-$valued $\mathcal{F}_t$ measurable random variable, $X_0^x = x$ P a.s, and $t \to X_t^x$ is *cad-lag* (i.e right-continuous with left-hand limits);

(ii) For each $f \in \mathcal{M}_b(M)$ the mapping

$$(t,x) \in \mathbb{R}^+ \times M \to P_t f(x) = \mathsf{E}(f(X_t^x)) \tag{3}$$

is measurable, and

$$\mathsf{E}\left[f(X_{t+s}^x)|\mathcal{F}_t\right] = (P_s f)(X_t^x),\ \mathsf{P}\ a.s. \tag{4}$$

Equation (3) defines a semigroup $(P_t)_{t\geq 0}$ of contractions on $\mathcal{M}_b(M)$. That is $P_t \circ P_s f = P_{t+s} f$ and $\|P_t f\| \leq \|f\|$.

We sometimes let $\mathbb{P}_x$ denote the law of (X_t^x) on the Skorokhod space $D(\mathbb{R}_+, M)$. That is $\mathbb{P}_x(\cdot) = \mathsf{P}(\omega \in \Omega\ : (X_t^x(\omega))_{t\geq 0} \in \cdot)$.
Our main assumption is the following:

Hypothesis 1 (Standing assumption) There exists a closed set $M_0 \subset M$ called the *extinction set* of $(P_t)_{t\geq 0}$ which is invariant under $(P_t)_{t\geq 0}$:

$$\forall t \geq 0 \; P_t \mathbf{1}_{M_0} = \mathbf{1}_{M_0}.$$

We let $M_+ = M \setminus M_0$ denote the *non extinction set*. Note that M_+ is open and invariant (i.e $P_t \mathbf{1}_{M_+} = \mathbf{1}_{M_+}$).

In addition to Hypothesis 1 we make certain regularity and tightness assumptions (Hypotheses 2 and 3 below) that will be needed throughout.

Hypothesis 2 ($C_b(M)$-Feller continuity) *For each $f \in C_b(M)$ the mapping $(t, x) \in \mathbb{R}_+ \times M \to P_t f(x)$ is continuous.*

Remark 1 For further reference we will call such a semigroup a $C_b(M)$-*Feller Markov semigroup*. This terminology is chosen to avoid confusion with the usual definition of Feller Markov semigroups (see e.g [25] or [51]) which assumes that P_t maps $C_0(M)$ into itself and induces a strongly continuous semigroup on $C_0(M)$. Note that every Feller semigroup is $C_b(M)$-Feller. When M is compact, all the examples considered here are Feller (in the usual sense). However, ecological stochastic differential equations on non compact spaces are usually not, as shown in the next example.

Example 1 (Logistic SDE) Consider the *logistic stochastic differential equation* on $\mathbb{R}^+$

$$dx = x((1-x)dt + \sigma dB_t).$$

Then, for all $t > 0$,

$$X_t^x = \frac{x e^{(1-\frac{\sigma^2}{2})t + \sigma B_t}}{1 + x \int_0^t e^{(1-\frac{\sigma^2}{2})s + \sigma B_s} ds} \to X_t^\infty := \frac{e^{(1-\frac{\sigma^2}{2})t + \sigma B_t}}{\int_0^t e^{(1-\frac{\sigma^2}{2})s + \sigma B_s} ds}$$

as $x \to \infty$. It easily follows that the induced semigroup doesn't preserve $C_0(\mathbb{R}^+)$ nor that it is strongly continuous on $C_b(\mathbb{R}^+)$. However, it is a $C_b(\mathbb{R}^+)$ Feller Markov semigroup.

Remark 2 The cad-lag continuity of the paths and Hypothesis 2 make (X_t^x) a strong Markov process (see e.g Theorem 6.17 in [51] stated for Feller (in the usual sense) Markov processes but the proof only requires cad-lag continuity and $C_b(M)$ Feller continuity).

We let $\mathcal{L}$ denote the *generator* of $(P_t)_{t\geq 0}$ on $C_b(M)$ and $\mathcal{D}(\mathcal{L}) \subset C_b(M)$ its *domain.* Here, following [64] (see also [63]) $\mathcal{D}(\mathcal{L})$ is defined as the set of $f \in C_b(M)$ for which

(i) $\mathcal{L}f(x) := \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t}$ exists for all $x \in M$;

(ii) $\mathcal{L}f \in C_b(M)$;

(iii) $\sup_{0<t\leq 1} \frac{1}{t}\|P_t f - f\| < \infty$.

It is easily seen that for all $f \in \mathcal{D}(\mathcal{L})$ and $t \geq 0$, $P_t f \in \mathcal{D}(\mathcal{L})$, and that for all $x \in M, t \mapsto P_t f(x)$ is C^1 and satisfies

$$\frac{d}{dt} P_t f(x) = \mathcal{L}(P_t f)(x) = P_t(\mathcal{L}f)(x). \tag{5}$$

Remark 3 In case (P_t) induces a strongly continuous semigroup on a Banach set $E \subset C_b(M)$, (for instance $C_b(M)$ or $C_0(M)$) the set $\{(f, g) \in E \times E\ f \in \mathcal{D}(\mathcal{L}), g = \mathcal{L}f\}$ equals the graph of the infinitesimal generator (defined in the usual sense) of (P_t) restricted to E.

Remark 4 For all $f \in C_b(M)$ and $\varepsilon > 0$ let $f_\varepsilon = \frac{1}{\varepsilon}\int_0^\varepsilon P_s f ds$. Then $f_\varepsilon \in \mathcal{D}(\mathcal{L}), \mathcal{L}(f_\varepsilon) = \frac{1}{\varepsilon}(P_\varepsilon f - f)$ and $\lim_{\varepsilon\to 0} f_\varepsilon = f$ (pointwise).

We let $\mathcal{D}^2(\mathcal{L})$ denote the set of $f \in C_b(M)$ such that both f and f^2 lie in $\mathcal{D}(\mathcal{L})$. If $f \in \mathcal{D}^2(\mathcal{L})$ we let

$$\Gamma(f) = \mathcal{L}f^2 - 2f\mathcal{L}f \tag{6}$$

denote the *carré du champ* of f. Note that $\Gamma(f) = \lim_{t\to 0} \frac{1}{t}(P_t f^2 - (P_t f)^2)$ so that $\Gamma(f) \geq 0$.

2.1 Empirical, invariant and ergodic probabilities

We denote the sequence of *empirical occupation measures* $(\Pi_t^x)_{t\geq 0}$ of the process $(X_t^x)_{t\geq 0}$ as

$$\Pi_t^x(B) = \frac{1}{t}\int_0^t \mathbf{1}_{\{X_s^x \in B\}} ds,\ \forall B \in \mathcal{B}(M). \tag{7}$$

Hence, $\Pi_t^x(B)$ is the proportion of time spent by the process in B up to time t.

A probability measure $\mu \in \mathcal{P}(M)$ is called *stationary* or *invariant* if

$$\mu P_t = \mu$$

for all $t \geq 0$, or equivalently, $\mu(P_t f) = \mu f$ for all $f \in \mathcal{M}_b(M)$ and all $t \geq 0$. We denote the set of invariant probability measures of $(P_t)_{t\geq 0}$ by $\mathcal{P}_{inv}(M)$. We also let

$$\mathcal{P}_{inv}(M_0) = \{\mu \in \mathcal{P}_{inv}(M) : \ \mu(M_0) = 1\},$$

and

$$\mathcal{P}_{inv}(M_+) = \{\mu \in \mathcal{P}_{inv}(M) : \ \mu(M_+) = 1\}.$$

A set $B \in \mathcal{B}(M)$ is called *invariant* if $P_t \mathbf{1}_B = \mathbf{1}_B$ for all $t \geq 0$. Invariant probability $\mu \in \mathcal{P}_{inv}(M)$ is called *ergodic* if every invariant set has $\mu-$measure 0 or 1. Equivalently, μ is ergodic if and only if it is extremal, meaning that it cannot be written as a nontrivial convex combination $\mu = \epsilon\mu_1 + (1-\epsilon)\mu_0$ with $0 < \epsilon < 1$ of two other distinct invariant measures $\mu_0, \mu_1 \in \mathcal{P}_{inv}(M)$.

Given a set $S \subset M$ (typically M, M_+ or M_0) we denote by

$$\mathcal{P}_{erg}(S) = \{\mu \in \mathcal{P}_{inv}(M),\ \mu(S) = 1, \mu \text{ ergodic}\}$$

the set of ergodic probability measures on S.

In order to control the behavior of the process at infinity and to ensure the tightness of $(\Pi_t^x)_{t\geq 0}$ (when M is noncompact) we shall assume the existence of a convenient *Lyapunov function*.

Recall that a continuous map $W : M \mapsto \mathbb{R}$ is called *proper* provided $\{x \in M : W(x) \leq R\}$ is compact for all $R > 0$.

Hypothesis 3 There exist proper maps $W, \tilde{W} : M \mapsto \mathbb{R}_+$, and a continuous function $LW : M \mapsto \mathbb{R}$ enjoying the following properties:

(i) For every compact set $K \subset M$ there exists $W_K \in \mathcal{D}^2(\mathcal{L})$ such that

- **(a)** $W|_K = W_K$ and $\mathcal{L}(W_K)|_K = LW|_K$,
- **(b)** $\forall x \in M \ \sup\{P_t(\Gamma(W_K))(x) : \ t \geq 0, K \ \text{ compact }\} < \infty$

(ii) $LW \leq -\tilde{W} + C$ for some $C \geq 0$.

Remark 5 If M is compact, Hypothesis 3 is automatically satisfied, say with $W = LW = \tilde{W} = 0$.

The next result ensures that, under Hypotheses 2 and 3, the empirical occupation measures (Π_t^x) is almost surely relatively compact and that its limit points are invariant. The proof is given in the appendix Section 9.1. Note that some versions of this results (for stochastic differential equations) are already proved in [70] and [26].

Theorem 2.1 *Assumes Hypotheses 2 and 3. Then*

(i) *For all $x \in M$*

$$0 \leq P_t W(x) + \int_0^t P_s(\tilde{W})(x) ds \leq W(x) + Ct.$$

(ii) *For all $x \in M$, P almost surely, $\limsup_{t\to\infty} \Pi_t^x \tilde{W} \leq C$, (Π_t^x) is tight, and every limit point of $(\Pi_t^x)_{t\geq 0}$ lies in $\mathcal{P}_{inv}(M)$. Furthermore $\mathcal{P}_{inv}(M)$ is compact and $\mu\tilde{W} \leq C$ for all $\mu \in \mathcal{P}_{inv}(M)$.*

(iii) *In case $\tilde{W} = \alpha W$ for some $\alpha > 0$,*

$$P_t W \leq e^{-\alpha t}(W - C/\alpha) + C/\alpha.$$

Remark 6 Note that, while (by Theorem 2.1) both $\mathcal{P}_{inv}(M)$ and $\mathcal{P}_{inv}(M_0)$ are non-empty, $\mathcal{P}_{inv}(M+)$ may be empty.

3 Motivating Examples

3.1 Pure jump ecological processes

The simplest examples are given by pure jump processes.

Let $M = \mathbb{R}_+^n = \{x \in \mathbb{R}^n : x_i \geq 0\}, (E, \mathcal{E}, \nu)$ a probability space (representing the *environment*) and for each $i = 1, \ldots, n$, $R_i : M \times E \mapsto \mathbb{R}_+^*$ a positive measurable mapping, continuous in the first variable.

Vector $x = (x_1, \ldots, x_n) \in M$ represents the state (abundances) of n interacting species and $R_i(x, e)$ the *fitness* of population i in environment e.

Let $(e_k)_{k\geq 1}$ be a sequence of i.i.d random variables distributed according to ν, and $(Y_k)_{k\geq 1}$ a discrete time Markov chain defined by

$$Y_{k+1} = G(Y_k, e_{k+1})$$

where

$$G(x,e) = (x_1R_1(x,e),\dots,x_nR_n(x,e)).$$

Such discrete time models of interacting populations in a fluctuating environment are analyzed in [70].

Let now (N_t) be a Poisson process with parameter $\lambda > 0$, independent of (e_k). The process

$$X_t = Y_{N_t}$$

is a jump Markov process on M. The associated semigroup is strongly continuous on $C_b(M)$ (as well as on $\mathcal{M}_b(M)$) and writes $P_tf = e^{t\mathcal{L}}f$ where $\mathcal{L}$ is the bounded operator on $C_b(M)$ defined by

$$\mathcal{L}f(x) = \int (f(G(x,e)) - f(x))\nu(de).$$

Here $\mathcal{D}(\mathcal{L}) = \mathcal{D}^2(\mathcal{L}) = C_b(M)$ and

$$\Gamma(f)(x) = \int [f(G(x,e)) - f(x)]^2\nu(de).$$

For any given subset $I \subset \{1,\dots,n\}$, let

$$M_0^I = \{x \in M \ : \prod_{i\in I} x_i = 0\} \tag{8}$$

be the set corresponding to the extinction of at least one of the species $i \in I$. Hypothesis 1 is clearly satisfied with $M_0 = M_0^I$. Hypothesis 2 is satisfied by strong continuity of (P_t). A sufficient condition ensuring Hypothesis 3 is given by the existence of suitable continuous Lyapunov function $W, \tilde{W} : M \mapsto \mathbb{R}_+$ such that $\lim_{\|x\|\to\infty} W(x) = \infty, \lim_{\|x\|\to\infty} \tilde{W}(x) = \infty$ and

$$W(G(x,e)) - W(x) \leq -\tilde{W}(x) + C, \ \forall e \in \mathcal{E}.$$

3.2 Ecological SDEs

Consider a stochastic differential equation on $M = \mathbb{R}_+^n$, having the form

$$dx_i = x_i^{\alpha_i}[F_i(x)dt + \sum_{j=1}^m \Sigma_i^j(x)dB_t^j], i = 1,\dots,n. \tag{9}$$

where F_i, Σ_i^j are real valued localy Lipschitz maps, Σ_i^j is bounded[1] $(B_t^1, \ldots, B_t^m)$ is an m-dimensional standard Brownian motion, and $\alpha_i \in \mathbb{N}^*$ (the set of positive integer).

The variables $(x_1, \ldots, x_n)$ typically represent the abundance of n interacting species, F_i is the per-capita growth rate of species i in absence of noise, and $(B_t^1, \ldots, B_t^m)$ models the environmental noise.

This type of process includes Brownian perturbations of Lotka-Volterra processes as considered in [27] as well as general stochastic ecological equation that have been recently considered by Hening and Nguyen in [37].

For any given subset $I \subset \{1, \ldots, n\}$, we let

$$M_0^I = \{x \in M \ : \prod_{i \in I} x_i = 0\} \tag{10}$$

denote the *extinction set* corresponding to the extinction of at least one of the species $i \in I$.

Remark 7 Abiotic or feedback variables can easily be included in the model by relaxing the assumption that α_i is nonzero. In this case the species abundances are the x_i for which $\alpha_i \neq 0$, the others x_i are the abiotic variables, and the state space is $\{x \in \mathbb{R}^n : \ \alpha_i x_i \geq 0\}$.

We let $a(x)$ denote the positive semi definite matrix defined by

$$a_{ij}(x) = \sum_{k=1}^m \Sigma_i^k(x) \Sigma_j^k(x). \tag{11}$$

For all $f : \mathbb{R}_+^n \mapsto \mathbb{R}, C^2$, we let

$$Lf(x) = \sum_i x_i^{\alpha_i} F_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j} x_i^{\alpha_i} x_j^{\alpha_j} a_{ij}(x) \frac{\partial^2 f}{\partial x_i x_j}(x) \tag{12}$$

and

$$\Gamma_L(f)(x) = \sum_{i,j} x_i^{\alpha_i} x_j^{\alpha_j} a_{ij}(x) \frac{\partial f}{\partial x_i}(x) \frac{\partial f}{\partial x_j}(x). \tag{13}$$

The next proposition gives conditions ensuring that hypotheses 1, 2, 3 hold. Its proof uses standard arguments given, for completeness, in appendix Section 9.2.

Recall that the maps F and Σ^j are locally Lipschitz with Σ^j bounded.

[1] This assumption is chosen here for simplicity and can be relaxed under other conditions as shown in [37].

Proposition 3.1 *Assume that there exist a C^2 proper[2] map $U : M \mapsto [1,\infty[$, a continuous function $\varphi : M \mapsto \mathbb{R}_+$, and constants $\alpha > 0, \beta \geq 0$ and $0 \leq \eta < 1$ such that*

$$LU \leq -\alpha U(1+\varphi) + \beta,$$

and

$$\Gamma_L(U) \leq cst(U^{2+\eta}).$$

Then

(i) *For each $x \in M$ there exists a unique (strong) solution $(X_t^x)_{t\geq 0} \subset M$ to (9) with initial condition $X_0^x = x$ and X_t^x is continuous in (t,x); In particular Hypothesis 2 holds.*

(ii) $\sup_{t\geq 0} \mathsf{E}(U(X_t^x)) \leq cst(1+U(x))$.

(iii) *Let $C_c^2(M)$ be the set of C^2 maps $f : M \mapsto \mathbb{R}$ with compact support[3]. Then $C_c^2(M) \subset \mathcal{D}^2(\mathcal{L})$ and for all $f \in C_c^2(M)$*

$$\mathcal{L}f(x) = Lf(x) \text{ and } \Gamma(f)(x) = \Gamma_L(f)(x).$$

(iv) *Hypothesis 1 holds true with $M_0 := M_0^I$*

(v) *Hypothesis 3 holds with*

$$W = U^{\frac{1-\eta}{2}}$$

and $\tilde{W} = (1+cst)W(1+\varphi)$.

Remark 8 (The Hening Nguyen condition) Set $\tilde{U} = \log(U)$. Then

$$LU = e^{\tilde{U}}(L\tilde{U} + \frac{1}{2}\Gamma_L(\tilde{U})) \text{ and } \Gamma_L(\tilde{U}) = \frac{1}{U^2}\Gamma_L(U)$$

so that the above conditions on U are equivalent to the conditions

$$\limsup_{\|x\|\to\infty} L\tilde{U} + \frac{1}{2}\Gamma_L(\tilde{U}) + \alpha(1+\varphi) < 0 \tag{14}$$

for some $\alpha > 0$ and

$$\Gamma_L(\tilde{U}) \leq cst(\exp \eta\tilde{U}) \tag{15}$$

---

[2] i.e $\lim_{\|x\|\to\infty} U(x) = \infty$

[3] By this we mean that is the restriction to M of a C^2 function $f : \mathbb{R}^n \mapsto \mathbb{R}$ with compact support.

for $\eta \geq 0$.

In particular, if $\hat{U} \geq 0$ is any C^2 proper function such that

$$\limsup_{\|x\| \to \infty} L(\hat{U}) + \alpha(1+\varphi) < 0 \tag{16}$$

and

$$\Gamma_L(\hat{U}) \leq cst \tag{17}$$

Then the conditions ((14), (15)) are satisfied for $\tilde{U} = \theta \hat{U}$ (i.e $U = e^{\theta \hat{U}}$), θ small enough and α replaced by $\alpha\theta$.

In case $\hat{U}(x) = \log(1 + \sum_i c_i x_i)$ with $c_i > 0$, this condition is the one assumed in [37].

Example 2 (Competitive Lotka-Volterra systems) Consider the general model given by (1) under the assumptions that $\alpha_i = 1$ and

$$F_i(x) \leq f_i(x_i)$$

where $f_i : \mathbb{R} \mapsto \mathbb{R}$ is continuous and

$$x_i > R \Rightarrow p(f_i(x_i) + \frac{(p-1)}{2} a_{ii}(x)) < -\alpha \tag{18}$$

for some positive numbers R, α and $p \geq 1$. Then the conditions of Proposition 3.1 are satisfied with $U(x) = 1 + \sum_i x_i^p, \varphi = 0$ and $\eta = 0$ (the verification is easy and left to the reader). A particular case is given by the class of *competitive Lotka-Volterra* systems for which

$$F_i(x) = r_i - \sum_j b_{ij} x_j \tag{19}$$

with $b_{ij} \geq 0$ and $b_{ii} > 0$. Here, it suffices to chose $f_i(x) = r_i - b_{ii} x_i$. Other examples include Lotka-Volterra mutualism systems as considered in [35].

Remark 9 (Ecological SDEs on the simplex) In numerous models occurring in ecology, population dynamics and game theory, x_i represents the proportion of species i rather that its abundance. The state space is then the unit simplex

$$\Delta^{n-1} = \{x \in \mathbb{R}^n_+ \, : \sum_i x_i = 1\}.$$

In this case, to insure invariance of Δ^{n-1} by (9), one assumes that the drift and diffusion vector fields are tangent to Δ^{n-1}. That is

$$\sum_{i=1}^{n} x_i^{\alpha_i} F_i(x) = \sum_{i=1}^{n} x_i^{\alpha_i} \Sigma_i^j(x) = 0.$$

Under these conditions, the processes (9) induces a Feller (in the usual sense) Markov process on a compact metric space, $M = \Delta^{n-1}$. In particular, Hypotheses 2 and 3 hold, while Hypothesis 1 obviously holds with M_0 defined by (10).

Such model have been considered by Foster and Young [28], Fudenberg and Harris [30], Hofbauer and Imhof [43]. A first general analysis of their persistence was first given by Benaim *et al.* [6] and generalized in Schreiber *et al.* [70].

3.3 Random ecological ODEs

Let $\{G^j\}_{j=1,\ldots,m}$ be a family of m vector fields on $\mathbb{R}_+^n$ having the form

$$G_i^j(x) = x_i^{\alpha_i} F_i^j(x);\ i = 1, \ldots n$$

where $\alpha_i \in \mathbb{N}^*$ and F_i^j is C^1.

Let $\Phi^j = \{\Phi_t^j\}$ be the local flow on $\mathbb{R}_+^n$ induced by the ordinary differential equation $\dot{x} = G^j(x)$.

We assume here for simplicity that there exists a compact set $B \subset \mathbb{R}_+^n$ positively invariant under each Φ^j. That is $\Phi_t^j(B) \subset B$ for all $t \geq 0$.

Let

$$M = B \times \{1, \ldots, m\}.$$

For each $(x, j) \in M$, let $(X_t^{x,j} = (x(t), J(t)))_{t \geq 0}$ be the process on M starting from (x, j) (i.e $X_0^{x,j} = (x, j)$) defined by

$$\begin{cases} \dfrac{dx(t)}{dt} = G^{J(t)}(x(t)), \\ \\ \mathsf{P}(J(t+s) = k | \mathcal{F}_t, J(t) = j) = a_{jk}(x(t))s + o(s) \end{cases} \tag{20}$$

where $\forall\ j, k \in \{1, \ldots m\}$, $a_{jk} : \mathbb{R}_+^n \mapsto \mathbb{R}_+$ is continuous nonnegative, $a_{jj} = 0$, and $\{a_{jk}(x)\}_{j,k}$ is irreducible for all x.

This type of process belongs to the larger class of *Piecewise deterministic Markov Processes*, a term coined by Davis [21]. Their ergodic properties have recently been the focus of much attention in the literature ([3], [8], [9], [20], [2], [5], [1], [7]). We refer the reader to the recent overview by Malrieu [54].

Let $C^1(M)$ be the set of maps $f : M \mapsto \mathbb{R}, (x,j) \mapsto f(x,j)$ which are C^1 in the x variable. It follows from Proposition 2.1 in [9] that $(X_t^{x,j})_{t\geq 0}$ is Feller, $C^1(M) \subset \mathcal{D}^2(\mathcal{L})$ and for all $f \in C^1(M)$

$$\mathcal{L}f(x,j) = \langle \nabla_x f(x,j), G^j(x)\rangle + \sum_{k=1}^{m} a_{jk}(x)(f(x,k) - f(x,j))$$

and

$$\Gamma(f)(x,j) = \sum_{k=1}^{m} a_{jk}(x)(f(x,k) - f(x,j))^2.$$

Let $I \subset \{1,\ldots,n\}$. Then Hypothesis 1 holds true with

$$M_0 = M_0^I = \{(x,j) \in M \ : \prod_{i\in I} x_i = 0\}.$$

4 Stochastic Persistence

The following definition, inspired by the seminal work of Chesson [15, 16], follows from Schreiber [69].

Definition 4.1 The family $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ is called *stochastically persistent (with respect to M_0)* if for all $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset M_+$ such that for all $x \in M_+$:

$$\mathsf{P}(\liminf_{t\to\infty} \Pi_t^x(K_\varepsilon) \geq 1 - \varepsilon) = 1. \tag{21}$$

If M_0 is unambiguous we simply say that $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ is stochastically persistent.

In models of population dynamics, the interpretation of stochastic persistence is that all the species, initially present, persist (stay away from the extinction set) over arbitrary long periods of time.

Remark 10 Suppose that $M_0 = M_0^1 \cup M_0^2$ where $M_0^{1,2}$ are closed and invariant under $(P_t)_{t\geq 0}$. If the process if stochastically persistent with respect to M_0^1 and M_0^2 then it is stochastically persistent with respect to M_0. Note that, however, the converse is false, as shown by the following deterministic example

Example 3 Consider the Rosenzweig MacArthur [66] prey predator model

$$\begin{cases} \frac{dx_1}{dt} = x_1(1 - \frac{x_1}{\kappa} - \frac{x_2}{1+x_1}) \\ \frac{dx_2}{dt} = x_2(-\alpha + \frac{x_1}{1+x_1}) \end{cases} \tag{22}$$

on the state space $M = \mathbb{R}_+^2$ where α, κ are positive parameter. Set $M_0^1 = \mathbb{R}_+ \times \{0\}, M_0^2 = \{0\} \times \mathbb{R}_+$ and $M_0 = M_0^1 \cup M_0^2$. Every trajectory on M_0^2 converges to the origin, so that the system is never persistent with respect to M_0^1. Assume $\alpha < \frac{\kappa}{1+\kappa}$. Then (see e.g [72]) the system admits an equilibrium $p \in M_+$. If $\alpha < \frac{\kappa-1}{\kappa+1}$ p is a source and there is and a limit cycle $\gamma \subset M_+$ surrounding p whose basin is $M_+ \setminus \{p\}$. If $\alpha \geq \frac{\kappa-1}{\kappa+1}$ every positive trajectory converges to p. This makes the system persistent with respect to M_0.

Proving or disproving stochastic persistence requires to control the behavior of (X_t) near the extinction set. This will be done by assuming the existence of another suitable type Lyapunov function.

Hypothesis 4 *There exist continuous maps $V : M_+ \mapsto \mathbb{R}_+$ and $H : M \mapsto \mathbb{R}$ enjoying the following properties:*

(i) *For all compact $K \subset M_+$, there exists $V_K \in \mathcal{D}^2(\mathcal{L})$ such that*

- **(a)** $V_K|_K = V|_K$ *and* $(\mathcal{L}V_K)|_K = H|_K$.
- **(b)** $\forall x \in M \sup\{P_t(\Gamma(V_K))(x) \ : K \subset M_+, \ K \text{compact}, t \geq 0\} < \infty$ *where Γ is defined by (6).*

(ii) *The map $\frac{\tilde{W}}{1+|H|}$ is proper, where $\tilde{W}$ is like in Hypothesis 3.*

We will sometimes assume the stronger version of **(ii)**:

(ii)' $|H|^q \leq cst(1 + \tilde{W})$ for some $q > 1$.

Note that by condition (ii) above and Theorem 2.1, $H \in L^1(\mu)$ for all $\mu \in \mathcal{P}_{inv}(M)$. The following definition then makes sense.

Definition 4.2 (H-exponents) If V and H are like in Hypothesis 4 we let

$$\Lambda^-(H) = -\sup\{\mu H \ : \mu \in \mathcal{P}_{erg}(M_0)\},$$

and

$$\Lambda^+(H) = -\inf\{\mu H \ : \mu \in \mathcal{P}_{erg}(M_0)\}$$

denote the H-*exponents* of (X_t).

Remark 11 The key point here is that while H is defined on **all** M, V is defined **only** on M_+ and typically $V(x) \to \infty$ when $x \to M_0$.

Actually, if V can be defined **on all** M with condition (i) of Hypothesis 4 valid **for all** $K \subset M$ compact, then (see Remark 19)

$$\Lambda^-(H) = \Lambda^+(H) = 0.$$

Definition 4.3 We call $\{(X_t^x)_{t \geq 0} : x \in M_+\}$ *H-persistent* if there exists (V, H) like in Hypothesis 4 such that $\Lambda^-(H) > 0$.

Example 4 (Logistic SDE, continuation of example 1) Consider the logistic equation given in Example 1. Here $M = \mathbb{R}_+$ and $M_0 = \{0\}$. Let $V :]0, \infty[\mapsto \mathbb{R}_+$ be any smooth function with bounded support (say, $V(x) = 0$ for $x \geq 1$) and coinciding with $-\log(x)$ on a neighborhood of 0. Then the map $H(x) = V'(x)x(1-x) + \frac{\sigma^2}{2}x^2V''(x)$ extends continuously to $\mathbb{R}_+$ and coincide with $x - 1 + \frac{\sigma^2}{2}$ on a neighborhood of 0. Clearly V and H satisfy Hypothesis 4, $\mathcal{P}_{erg}(M_0) = \{0\}$ and

$$\Lambda^+(H) = \Lambda^-(H) = 1 - \frac{\sigma^2}{2}.$$

Here $H-$ persistence simply writes

$$1 - \frac{\sigma^2}{2} > 0.$$

More sophisticated examples will be studied later.

Remark 12 By the ergodic decomposition theorem, compactness of $\mathcal{P}_{inv}(M_0)$ (Theorem 2.1) and continuity of $\mu \mapsto \mu H$ (Lemma 9.1 (ii)) combined with condition (ii) in Hypothesis 4, the following conditions are equivalent :

(a) $\Lambda^-(H) > 0$,

(b) $\mu H < 0$ for all $\mu \in \mathcal{P}_{erg}(M_0)$,

(c) $\mu H < 0$ for all $\mu \in \mathcal{P}_{inv}(M_0)$.

Similar to Remark 10 is the following

Remark 13 Suppose that $M_0 = M_0^1 \cup M_0^2$ where $M_0^{1,2}$ are closed and invariant under $(P_t)_{t\geq 0}$. Let $M_+^i = M \setminus M_0^i$, and $V^i : M_+^i \mapsto \mathbb{R}_+$, $H^i : M \mapsto \mathbb{R}$ be as in Hypothesis 4. Let V be defined on M_+ by $V = V^1 + V^2$ and let $H = H^1 + H^2$.

Then for all $\mu \in \mathcal{P}_{erg}(M_0)$ either

(i) $\mu(M_0^1) = 0$ and $\mu H = \mu H^2$, or

(ii) $\mu(M_0^2) = 0$ and $\mu H = \mu H^1$, or

(iii) $\mu(M_0^1 \cap M_0^2) = 1$ and $\mu H = \mu H^1 + \mu H^2$.

In particular if the process is H-persistent with respect to $M_0^i, i = 1, 2$ it is H-persistent with respect to M_0.

4.1 H-Persistence implies Stochastic Persistence

From now on, hypotheses 1 to 4 are implicitly assumed. The main result of this section is given by the following theorem whose proof is postponed to Section 7.

Theorem 4.4 *Assume that* $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ *is* H*-persistent. Then*

(i) *For all* $x \in M_+$*, every weak limit point of* $(\Pi_t^x)_{t\geq 0}$ *lies in* $\mathcal{P}_{inv}(M_+)$ *a.s.*

(ii) $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ *is stochastically persistent.*

This theorem has the following immediate consequence.

Corollary 4.5 *Assume that* $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ *is* H*-persistent and that* $\mathcal{P}_{inv}(M_+)$ *has cardinal at most one. Then,* $\mathcal{P}_{inv}(M_+)$ *has cardinal one, and letting* $\mathcal{P}_{inv}(M_+) = \{\Pi\}$*, for all* $x \in M_+$*,* $\Pi_t^x \Rightarrow \Pi$ *a.s. as* $t \to \infty$*.*

For further references, the probability Π in Corollary 4.5 is called the *persistent measure.* In ecological models, the persistence measure describes the long term behavior of coexisting species.

4.2 Support and Irreducibility

Unlike in deterministic models where persistence equates the existence of an attractor bounded away from the extinction set, the support of the persistent measure may well have nonempty intersection with M_0.

The nature of this support provides useful information on the dynamics. For specific models (see for instance [10] section 4, and [55] section 5) it can be computed by using some elementary control theory type arguments that we now briefly discuss. The general definitions given here will be rephrased in Sections 5 and 6 in terms of deterministic control systems.

Point $y \in M$ is said *accessible* from $x \in M$ if for every neighborhood U of y there exists $t \geq 0$ such that $P_t(x, U) = P_t \mathbf{1}_U(x) > 0$. We let $\mathbf{\Gamma}_x$ denote the set of points y that are accessible from x. For $A \subset M$ we let

$$\mathbf{\Gamma}_A = \cap_{x \in A} \mathbf{\Gamma}_x$$

denote the (possibly empty) closed set of points that are accessible from every $x \in A$.

Corollary 4.6 *Under the assumptions of Corollary 4.5,*

$$supp(\Pi) = \mathbf{\Gamma}_{M_+} = \mathbf{\Gamma}_x$$

for all $x \in \mathbf{\Gamma}_{M_+} \cap M_+$.

Proof: Let $O \subset M$ be an open set such that $\Pi(O) > 0$. Then, by Fatou lemma, Corollary 4.5 and Portmanteau theorem, for all $x \in M_+$

$$\liminf_{t \to \infty} \frac{1}{t} \int_0^t P_s \mathbf{1}_O(x) ds \geq \mathsf{E}(\liminf_{t \to \infty} \Pi_t^x(O)) \geq \Pi(O) > 0.$$

This proves that $supp(\Pi) \subset \mathbf{\Gamma}_{M_+}$.

Conversely, let $R(x, dy)$ be the resolvent kernel defined by $Rf = \int_0^\infty e^{-s} P_s f ds$. We claim that for every $y \in \mathbf{\Gamma}_{M_+}$, O a neighborhood of y and $x \in M_+$ $R(x, O) > 0$. By accessibility, there exists $t > 0$ such that $P_t(x, O) > 0$. Thus, by right continuity and Fatou Lemma

$$\liminf_{s \to t, s > t} P_s(x, O) \geq \mathsf{E}(\liminf_{s \to t, s > t} \mathbf{1}_O(X_s^x)) \geq P_t(x, O) > 0.$$

This proves that $s \to P_s(x, O)$ is positive on some interval $[t, t + \varepsilon]$, hence $R(x, O) > 0$.

Now, by invariance, $\Pi = \Pi R$. Therefore, $\Pi(O) = \int_{M_+} \Pi(dx) R(x, O) > 0$. This proves that $\mathbf{\Gamma}_{M_+} \subset supp(\Pi)$.

We now prove the last assertion. By definition $\mathbf{\Gamma}_{M_+} \subset \mathbf{\Gamma}_x$ for all $x \in M_+$. It then suffices to show that $\mathbf{\Gamma}_x \subset \mathbf{\Gamma}_{M_+}$ for $x \in \mathbf{\Gamma}_{M_+} \cap M_+$. Let $y \in \mathbf{\Gamma}_x$ and O a neighborhood of y. Then for some $t \geq 0$ and $\delta > 0$ $P_t(x, O) > \delta$. By $C_b(M)$ Feller continuity and Portmanteau Theorem, the set $V = \{z \in M_+ : P_t(z, O) > \delta\}$ is an open neighborhood of x. But since $x \in \mathbf{\Gamma}_{M_+}$ for all $z \in M_+$ there is some $s > 0$ such that $P_s(z, V) > 0$ Thus $P_{t+s}(z, O) \geq \delta P_s(z, V) > 0$.
$\square$

Remark 14 The preceding proof also shows that $\mathbf{\Gamma}_{M^+} \subset supp(\Pi)$ for all $\Pi \in \mathcal{P}_{inv}(M^+)$ even if $\mathcal{P}_{inv}(M^+)$ has cardinal greater than 1. However, in this case, $\mathbf{\Gamma}_{M^+}$ may be empty.

A sufficient (although non-necessary) condition ensuring that $\mathcal{P}_{inv}(M_+)$ has cardinal at most one (hence one when the process is stochastically persistent) is given by ψ-irreducibility, in the sense of Meyn and Tweedy (see [57] or [22]). A practical condition (implying ψ-irreducibility) is the existence of an accessible open *petite set*.

Proposition 4.7 *Assume there exist $x^* \in \mathbf{\Gamma}_{M_+}$ (in particular $\mathbf{\Gamma}_{M_+} \neq \emptyset$), a neighborhood U of x^*, a non zero measure ξ on M_+, and a probability measure γ on $\mathbb{R}_+$ such that for all $x \in U$*

$$R_\gamma(x, \cdot) := \int P_t(x, \cdot)\gamma(dt) \geq \xi(\cdot).$$

Then $\mathcal{P}_{inv}(M_+)$ has cardinal at most one. If furthermore, (X_t) is H-persistent, then $\mathcal{P}_{inv}(M_+) = \{\Pi\}, \Pi_t^x \Rightarrow \Pi$ a.s. for all $x \in M^+$; and

$$\lim_{t \to \infty} \Pi_t^x f = \Pi f$$

a.s. for all $f \in L^1(\Pi)$ and $x \in M^+$.

Proof: As shown in the proof of the last corollary, accessibility implies that for all $x \in M_+$ $R(x, U) > 0$, where R is the resolvent. Thus $RR_\gamma(x, \cdot) \geq \int_U R(x, dy) R_\gamma(y, \cdot) \geq R(x, U)\xi(\cdot)$. This proves that RR_γ is ξ irreducible on M_+. Thus, RR_γ, and therefore P_t, has at most one invariant probability on M_+ (see e.g [57] or [22]). If furthermore (X_t) is H persistent, then $\Pi_t^x \Rightarrow$

Π a.s by Corollary 4.5. It remains to prove the last assertion. The set U is, by assumption, a *petite set* in the sense of Meyn and Tweedy [57]. By Portmanteau theorem, accessibility and Corollary 4.6, $\liminf \Pi_t^x(U) \geq \Pi(U) > 0$ proving that U is recurrent. Now, the existence of a petite and recurrent set makes (X_t) *Harris recurrent* on M^+, and since $\mathcal{P}_{inv}(M_+)$ is nonempty (X_t) is positively recurrent on M^+. □

Remark 15 In many cases there exists a measure λ on M such that $\lambda P_t \ll \lambda$. A typical situation is when $M \subset \mathbb{R}^n$, λ the is the Lebesgue measure on $\mathbb{R}^n$ and $x \mapsto X_t^x(\omega)$ is a C^1 diffeomorphism for all (or P almost all) ω. If in addition, the assumptions of Proposition 4.7 are satisfied with $\xi \ll \lambda$, then

$$\Pi \ll \lambda.$$

Indeed, by Lebesgue decomposition Theorem $\Pi = \Pi_{ac} + \Pi_s$ with $\Pi_{ac} \ll \lambda$ and $\Pi_s \perp \lambda$. The proof of Proposition 4.7 easily implies that $\xi \ll \Pi$. Thus $\Pi_{ac} \neq 0$ because $\xi \ll \lambda$. By invariance $\Pi_{ac} + \Pi_s = \Pi_{ac} P_t + \Pi_s P_t$. Thus $\Pi_{ac} \geq \Pi_{ac} P_t$ by uniqueness of Lebesgue decomposition. This shows that $\frac{\Pi_{ac}}{\Pi_{ac}(M)}$ is excessive, hence invariant. That is $\Pi = \Pi_{ac}$.

4.3 Convergence

The next result shows that if the measure γ in Proposition 4.7 can be chosen to be a dirac mass then the law of X_t^x converges in total variation to Π whenever $x \in M_+$.

Recall that the *total variation distance* between two probabilities $\alpha, \beta \in \mathcal{P}(M)$ is defined as

$$|\alpha - \beta|_{TV} = \sup\{|\alpha(f) - \beta(f)| \ : \ f \in \mathcal{M}_b(M), \ \|f\|_\infty \leq 1\}.$$

We say that $x^* \in M$ is a *Doeblin point* if there exist a neighborhood U of x^*, a non zero measure ξ on M, and $t^* > 0$ such that for all $x \in U$

$$P_{t^*}(x, \cdot) \geq \xi(\cdot). \tag{23}$$

If a Doeblin point is accessible, the minorization condition (23) extends to every compact space. More precisely

Lemma 4.8 *Let $x^* \in \mathbf{\Gamma}_{M_+}$ be a Doeblin point and $x_0 \in supp(\xi) \cap M_+$ where ξ is like in (23). Then there exist a neighborhood $A \subset M_+$ of x_0, a probability ν on A (i.e $\nu(A) = 1$) and positive numbers T, c such that:*

(i) *For all $x \in A$ $P_T(x, \cdot) \geq c\nu(\cdot)$;*

(ii) *For every compact set $K \subset M_+$ there exist $n_K \in \mathbb{N}, c_K > 0$ such that $P_{n_k T}(x, \cdot) \geq c_K \nu(\cdot)$ for all $x \in K$;*

Proof: Let t^* and U be like in (23). By accessibility there exist $t_0, \delta > 0$ such that $P_{t_0}(x_0, U) > \delta$. By $C_b(M)$-Feller continuity (Hypothesis 2) and Portmanteau's theorem there exist $\varepsilon > 0$ and an open neighborhood A of x_0 such that $P_t(x, U) > \delta$ for all $x \in A$ and $|t - t_0| < \varepsilon$. Set $T = t_0 + t^*, \nu(\cdot) = \frac{\xi(\cdot \cap A)}{\xi(A)}$ and $c = \delta \xi(A)$. Then, $P_\tau(x, \cdot) \geq c\nu(\cdot)$ for $|\tau - T| < \varepsilon$. This proves (i).

Let $O(t, r) = \{x \in M_+ : P_t(x, U) > r\}$. The family $\{O(t, r)\}_{t \geq 0, r > 0}$ is an open (by $C_b(M)$-Feller continuity) covering (by accessibility) of M_+. Thus, for $K \subset M_+$ compact, $K \subset \cup_{i=1}^l O(t_i, r)$ for some $l \in \mathbb{N}, t_1, \ldots, t_l \geq 0$ and $r > 0$. Choose k large enough so that $\frac{t_i + t^*}{k} < \varepsilon$ for all $i = 1, \ldots, l$ and set $\tau_i = T - \frac{t_i + t^*}{k}$. Then for $x \in O(t_i, r)$

$$P_{kT}(x, \cdot) = P_{t_i + t^* + k\tau_i}(x, A) \geq r\xi(A)(c\nu(A))^{k-1}\nu(\cdot).$$

This proves (ii). $\square$

Theorem 4.9 *Assume that $\{(X_t^x)_{t \geq 0} : x \in M_+\}$ is H-persistent and that there exists a Doeblin point $x^* \in \mathbf{\Gamma}_{M_+}$. Then*

$$\mathcal{P}_{inv}(M_+) = \{\Pi\}$$

and for all $x \in M_+$

$$\lim_{t \to \infty} |P_t(x, \cdot) - \Pi|_{TV} = 0.$$

Proof: We use the notation of Lemma 4.8. Let Y be the discrete chain on M_+ whose transition kernel is P_T (restricted to M_+). By Lemma 4.8 (i), A is a *small* set for Y and, by (ii), it is accessible for Y from every point in M_+. In addition, by Proposition 4.7, Y has an invariant probability implying that A is *recurrent*. By application of Orey's theorem (see e.g Theorem 8.3.18 in [22]), the existence of a *small accessible recurrent set* imply that

$$\lim_{n \to \infty} |\mu P_T^n - \Pi|_{TV} = 0$$

for all $\mu \in \mathcal{P}(M_+)$. Now, writing $t = n_t T + r_t$ with $0 \leq r_t < T, n_t \in \mathbb{N}$,

$$\lim_{t\to\infty} |\mu P_t - \Pi|_{TV} = \lim_{t\to\infty} |\mu P_{n_t T} P_{r_t} - \Pi P_{r_t}|_{TV} \leq \lim_{n\to\infty} |\mu P_{nT} - \Pi|_{TV} = 0.$$

□

4.4 Rate of Convergence

Under certain additional assumptions, the rate of convergence in Theorem 4.9 can be shown to be exponential.

Throughout this section we will assume the following strengthening of Hypothesis 4:

Hypothesis 5 (strong version of Hypothesis 4) *V and H are like in Hypothesis 4, and in addition:*

(a) *The jumps of $V(X_s)$ are almost surely bounded. That is*

$$|V(X_s) - V(X_{s-})| \leq \Delta V,$$

where $0 \leq \Delta V < \infty$.

(b) *Condition $(i), (b)$ of Hypothesis 4 is strengthened to*

$$\gamma := \sup\{\|\Gamma(V_K)\| \ : K \subset M_+, K \ \textit{compact}\ \} < \infty$$

For further reference, we will say that $\{(X_t^x)_{t\geq 0}\ x \in M_+\}$ is *H-persistent, strong version* if it satisfies Hypothesis 5 and is H-persistent. If additionally condition $(ii)'$ in Hypothesis 4 is verified, we will say that it is *H-persistent, strong version'*.

The case M_0 compact

We first consider the situation where M_0 is compact. We let

$$M_0^\delta = \{x \in M_+ \ : d(x, M_0) < \delta\}$$

denote the δ neighborhood of M_0.

Theorem 4.10 *Assume that* $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ *is* H*-persistent (strong version),* M_0 *is compact,* $\tilde{W} = \alpha W$ *for some* $\alpha > 0$ *(where* W *and* $\tilde{W}$ *are like in Hypothesis 3) and that there exists a Doeblin point* $x^* \in \mathbf{\Gamma}_{M_+}$*. Then, there exist* $\lambda, \theta > 0, cst$ *such that for all* $x \in M_+$ *and* $f : M^+ \mapsto \mathbb{R}$ *measurable,*

$$|P_t f(x) - \Pi f| \leq cst(1 + W_\theta(x)) e^{-\lambda t} \|f\|_{W_\theta},$$

where W_θ *is continuous, lies in* $L^1(\Pi)$*, and coincide with* $V_\theta = e^{\theta V}$ *on* M_0^δ *and with* W *on* $\{W > R\}$ *for some* $R > 0$*. Here*

$$\|f\|_{W_\theta} = \sup_{x \in M^+} \frac{|f(x)|}{1 + W_\theta(x)}.$$

In particular,

$$|P_t(x, \cdot) - \Pi|_{TV} \leq cst(1 + W_\theta(x)) e^{-\lambda t}$$

Proof of Theorem 4.10. The following lemma follows from Proposition 8.2 proved in Section 8.

Lemma 4.11 *There exist positive numbers* $\theta, \kappa, T_0 < T_1$, $0 < \rho < 1$, *and a continuous function* $V_\theta : M_+ \mapsto \mathbb{R}_+$ *such that*

(i) $V_\theta = e^{\theta V}$ *on* M_0^δ *for some* $\delta > 0$.

(ii) $\lim_{x \to M_0} V_\theta(x) = \infty$ *and* V_θ *is bounded on* $M_+ \setminus M_0^\delta$;

(iii) *For all* $T \in [T_0, T_1]$, $P_T V_\theta \leq \rho V_\theta + \kappa$.

The proof of Theorem 4.10 is now a consequence of a classical result often refereed as "Harris's theorem" which proof can be found in numerous places (e.g [22], [57]). Here we rely on the following version given (and proved) by Hairer and Mattingly [36]. Let $\mathcal{P}$ be a Markov kernel on a measurable space E. Assume that:

(i) There exists a map $\mathcal{W} : E \mapsto [0, \infty[$ and constants $0 < \gamma < 1, \tilde{K} \geq 0$ such that $\mathcal{P}\mathcal{W} \leq \gamma \mathcal{W} + \tilde{K}$

(ii) For some $R > \frac{2\tilde{K}}{1-\gamma}$ there exists a probability measure ν and a constant c such that $\mathcal{P}(x, .) \geq c\nu(.)$ whenever $\mathcal{W}(x) \leq R$.

Then there exists a unique invariant probability Π for $\mathcal{P}$, and constants $0 \leq \tilde{\gamma} < 1, cst \geq 0$ such that for every measurable map $f : E \mapsto \mathbb{R}$ and all $x \in E$

$$|\mathcal{P}^n f(x) - \Pi f| \leq cst\, \tilde{\gamma}^n (1 + \mathcal{W}(x)) \|f\|_{\mathcal{W}}.$$

Here $\|f\|_{\mathcal{W}} = \sup_{x \in E} \frac{|f(x)|}{1+\mathcal{W}(x)}$. To apply this result, set $E = M_+$ and, using the notation of Lemma 4.11, $\mathcal{W} = V_\theta + W$. Proposition 4.11 combined with the fact that $P_t W \leq e^{-\alpha t} W + C/\alpha$ (see Theorem 2.1 (iii)) yield

$$P_{nT} \mathcal{W} \leq \tilde{\rho}^n \mathcal{W} + \tilde{K} \tag{24}$$

for all $n \in \mathbb{N}$ and $T_0 \leq T \leq T_1$, with $\tilde{\rho} = \max\{\rho, e^{-\alpha T_0}\}$ and $\tilde{K} = \frac{\kappa}{1-\rho} + \frac{C}{\alpha}$.

Choose $R > \frac{2\tilde{K}}{(1-\tilde{\rho})^2}$. The set $\mathcal{W}_R = \{x \in M_+ : \mathcal{W} \leq R\}$ is a compact subset of M_+ and by Lemma 4.8, there exist some constants $T_R, c_R > 0$ depending on R and a probability measure ν on M_+ - which we can assume to be supported by $\mathcal{W}_R$- such that $P_{T_R}(x, .) \geq c_R \nu(.)$ for all $x \in \mathcal{W}_R$. By iteration, this gives $P_{mT_R}(x, .) \geq c_R^m \nu(.)$ for all $x \in \mathcal{W}_R$ and $m \in \mathbb{N}^*$. Choose now $T \in [T_0, T_1]$ such that T_R/T is rational, and positives integers m, n such that $m/n = T_R/T$. Thus $P_{nT} = P_{mT_R} = \mathcal{P}$ verifies conditions $(i), (ii)$ above of Harris's theorem with $\gamma = \tilde{\rho}^n$. The end of the proof is similar to the end of the proof of Theorem 4.9.

□

The case M_0 noncompact

This section is strongly inspired by the recent beautiful work of Hening and Ngyuen [37] on Kolmogorov systems. When M_0 is noncompact, the existence of a Lyapunov function W controlling the behavior of the process at infinity, doesn't seem to be sufficient to ensure an exponential rate of convergence and one need to control the behavior of H at infinity.

We say that $\{(X_t^x)_{t \geq 0} : x \in M_+\}$ is *H-persistent with respect to M_0* **and** *at infinity*, it is H-persistent and the maps (V, H) of Definition 4.3 satisfy the two following additional properties:

(a) V is proper;

(b) There exists a compact $C \subset M$ such that

$$\sup_{x \in M \setminus C} H(x) < 0.$$

Theorem 4.12 *Assume that* $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ *is* H*-persistent (strong version') with respect to* M_0 *and at infinity and that there exists a Doeblin point* $x^* \in \mathbf{\Gamma}_{M_+}$. *Then there exists* $\lambda > 0, \theta, cst$ *such that for all* $x \in M_+$ *and* $f : M^+ \mapsto \mathbb{R}$ *measurable,*

$$|P_t f(x) - \Pi f| \leq cst(1 + W_\theta(x))e^{-\lambda t}\|f\|_{W_\theta}.$$

In particular,

$$|P_t(x, \cdot) - \Pi|_{TV} \leq cst(1 + W_\theta(x))e^{-\lambda t}$$

Here $W_\theta = e^{\theta V}$ *and* $\|f\|_{W_\theta}$ *is like in Theorem 4.10.*

The proof is given in section 8.1.

The "construction" of V (and H) ensuring persistence is (at least in all the examples we have in mind) dictated by our knowledge of the behavior of the process near the extinction set and there is, in general, no reason that the additional conditions (a) and (b) above ensuring persistence at infinity are equally valid. The following simple result is a useful trick to get around this problem.

Proposition 4.13 *Let* $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ *be* H*-persistent (strong version') with respect to* M_0. *Assume that there exist continuous functions* $(\tilde{V}, \tilde{H})$ *satisfying Hypothesis 5 with condition* $(ii)'$ *of Hypothesis 4 and such that:*

- *(i)* $\tilde{V}$ *is defined on* **all** M *and proper;*
- *(ii) Conditions* (i) *in Hypothesis 4 (respectively* (b) *in Hypothesis 5) are valid for* **every** *compact subset of* M;
- *(iii)* $\limsup_{x\to\infty} \varepsilon H(x) + \tilde{H}(x) < 0$, *for some* $\varepsilon > 0$.

Then the process is H*-persistent (strong version') at* M_0 *and at* ∞.

Proof: First assume that for all $K \subset M^+$ compact, $V_K + \tilde{V_K} \in \mathcal{D}^2(\mathcal{L})$. Then $(\varepsilon V + \tilde{V}, \varepsilon H + \tilde{H})$ satisfies Hypothesis 5 and condition $(ii)'$ of Hypothesis 4. This easily follows from the linearity of $\mathcal{L}$ and the property $\Gamma(f+g) \leq (\sqrt{\Gamma(f)} + \sqrt{\Gamma(g)})^2$ valid for $f, g \in \mathcal{D}^2(\mathcal{L})$ with $fg \in \mathcal{D}(\mathcal{L})$). By remark 11, $\Lambda(\varepsilon H + \tilde{H}) = \varepsilon\Lambda(H) > 0$. Hence the result.

In general (if we cannot argue that $V_K + \tilde{V_K} \in \mathcal{D}^2(\mathcal{L})$) note that (with the notation of Lemma 50) $M_t^V, M_t^{\tilde{V}}$ being square integrable martingales, the same is true for $M_t^{V+\tilde{V}} = M_t^V + M_t^{\tilde{V}}$ and $\langle M^{V+\tilde{V}}\rangle_t \leq (\sqrt{\langle M_t^V\rangle} + \sqrt{\langle M^{\tilde{V}}\rangle_t})^2$ and the proof goes through. $\square$

5 Application to Ecological SDEs (ii)

Consider the ecological SDE defined by (9). Let $I \subset \{1,\ldots,n\}$ and

$$M_0 := M_0^I = \{x \in M \ : \prod_{i\in I} x_i = 0\}$$

be the set corresponding to the extinction of at least one of the species $i \in I$.

Following [69], [70], [6], define the *invasion rate* of species i with respect to x as

$$\lambda_i(x) = F_i(x) - \alpha_i \frac{a_{ii}(x)}{2} x_i^{\alpha_i - 1} \tag{25}$$

and the *invasion rate* of species i with respect to $\mu \in \mathcal{P}_{erg}(M_0^I)$ as

$$\mu\lambda_i = \int \lambda_i d\mu \tag{26}$$

provided $\lambda_i \in L^1(\mu)$.

The following result asserts that if a weighted combination of the invasion rates $\{\mu\lambda_i\}_{i\in I}$ is positive for all $\mu \in \mathcal{P}_{erg}(M_0^I)$, then the process is H-(hence stochastically) persistent. This criterion goes back to the early work of Hofbauer [41] (see also [68] and [31]) but has been shown to apply also for SDEs, only recently, first in [6] (for small noise), then in [70] (on compact state spaces) and recently in [37] (on $\mathbb{R}_+^n$ for nondegenerate noise).

Note here that there is no assumption that the diffusion matrix (defined by (11)) is nondegenerate. We then retrieve Hofbauer's criterion, and - more importantly - this allows to handle situations where the "noise" only affect certain variables. Examples will be given in Section 5.1.

Theorem 5.1 *Let U, φ and η be as in Proposition 3.1. Assume that*

$$\limsup_{x\to\infty} \frac{U^{\frac{1-\eta}{2}}(x)(1+\varphi(x))}{1+\sum_{i\in I}|F_i(x)|} = \infty \tag{27}$$

and

$$\sum_{i=1}^n x_i^{\alpha_i-1} \leq cst\sqrt{U(x)} \tag{28}$$

Then

(i) *For all* $\mu \in \mathcal{P}_{erg}(M_0^I)$ *and* $i \in I$ $\lambda_i \in L^1(\mu)$ *and*

$$\mu\lambda_i \neq 0 \Rightarrow supp(\mu) \subset M_0^i = \{x \in M \ : x_i = 0\}.$$

(ii) *If there exist positive numbers* $\{p_i\}_{i\in I}$ *such that for all* $\mu \in \mathcal{P}_{erg}(M_0^I)$

$$\sum_{i\in I} p_i(\mu\lambda_i) > 0; \tag{29}$$

Then the process is H*-persistent with respect to* M_0^I.

(iii) *If furthermore* $\eta = 0, \alpha_i = 1$,

$$1 + \varphi + \varepsilon F_i \geq 0 \tag{30}$$

for some $\varepsilon > 0$ *and* $i \in I$, *and (27) is strengthened to*

$$\mid \frac{LU}{U} \mid^q + \sum_{i\in I} \mid F_i \mid^q \leq cst\sqrt{U} \tag{31}$$

for some $q > 1$. *Then, under condition (29), the process is* H*-persistent with respect to* M_0^I *(strong version') and persistent at infinity.*

Proof: (i) Condition (27) combined with Theorem 2.1 and Proposition 3.1 (v) imply that $\lambda_i \in L^1(\mu)$ for all $\mu \in \mathcal{P}_{inv}(M)$. The second assertion will be proved after the proof of assertion (ii).

(ii) For all $i \in I$ let $h_i(u) = \log(\frac{1}{u})$ if $\alpha_i = 1$, and $h_i(u) = \frac{u^{1-\alpha_i}}{\alpha_i - 1}$ if $\alpha_i > 1$. Let $v : \mathbb{R} \mapsto \mathbb{R}_+$ be a smooth function with bounded first and second derivatives such that $v(t) = t$ for $t \geq 1$. Set

$$V(x) = v(\sum_{i\in I} p_i h_i(x_i))$$

and

$$H(x) = v'(\sum_{i\in I} p_i h_i(x_i))(-\sum_i p_i\lambda_i(x)) + \frac{1}{2}v''(\sum_{i\in I} p_i h_i(x_i))\langle a(x)p, p\rangle \tag{32}$$

for $x \in M_+$. Then V (respectively H) coincide with $\sum_{i\in I} p_i h_i$ (respectively $\sum_{i\in I} -p_i\lambda_i$ on the set $\{x \in M_+ \ : \sum_{i\in I} p_i h_i(x) > 1\}$ and H extends continuously to M_0. Furthermore

$$|H| \leq cste(\sum_{i\in I} p_i|\lambda_i| + \sum_{i\in I} p_i^2)$$

so that, condition (27), imply that condition (ii) of Hypothesis 4 is satisfied. Let $b : \mathbb{R}_+ \mapsto [0,1]$ be a smooth function such that $b(t) = 1$ for $t \leq 1$ and $b(t) = 0$ for $t \geq 2$, and let $B(t) = \int_0^t b(u)du$. For all $n \geq 1$, set

$$V_n(x) = nB(\frac{V(x)}{n})b(\frac{\log(1+\sum_i x_i)}{n})$$

for $x \in M_+$ and

$$V_n(x) = nB(2)b(\frac{\log(1+\sum_i x_i)}{n})$$

for $x \in M_0$. Then $V_n \in \mathcal{D}^2(\mathcal{L})$ (because V_n is smooth with compact support), $V_n = V$ and $\mathcal{L}(V_n) = H$ on the set $K_n = \{x \in M_+ : \; V(x) \leq n, \log(1 + \sum_i x_i) \leq n\}$. Furthermore

$$\Gamma(V_n)(x) \leq cst(1 + \frac{\sum_i x_i^{\alpha_i}}{1+\sum_i x_i})^2 \leq cst(1+\sum_i x_i^{\alpha_i - 1})^2 \leq cst(1+U(x))$$

so that assumption (i) of Hypothesis 4 is satisfied in view of assertion (ii) of Proposition 3.1. This proves assertion (ii). Also, by Lemma 7.5, $\mu H > 0 \Rightarrow \mu(M_0^I) > 0$ which concludes the proof of (i).

(iii) When $\alpha_i = 1$, $\Gamma(V_n)(x) \leq cst$ and Hypothesis 5 is satisfied. By condition (31), condition $(ii)'$ in Hypothesis 4 also holds, so that the process is H-persistent (strong version').

We will now apply Proposition 4.13. Let $\tilde{V} = \log(U)$ and $\tilde{H} = L\tilde{U}$. Then

$$\tilde{H} = \frac{LU}{U} - \frac{1}{U^2}\Gamma_L(U) \leq -\alpha(1+\varphi) + \frac{\beta}{U}.$$

Using the right hand side equality, the assumptions on U, and condition (31), it is easily checked that $(\tilde{V}, \tilde{H})$ satisfies Hypothesis 5 and condition $(ii)'$ of Hypothesis 4. It suffices to set $\tilde{V}_n = nB(\frac{\tilde{V}}{n})$ and to argue as previously. From the left hand side inequality we get that

$$\limsup_{x\to\infty} \tilde{H} + \alpha(1+\varphi) \leq 0.$$

From condition (30) we get that

$$\sum_i p_i(1+\varphi) + \varepsilon \sum_i p_i F_i \geq 0.$$

Hence, replacing ε by a sufficiently smaller ε,

$$2\alpha(1+\varphi) - \varepsilon H \geq 0$$

where H is defined by (32). This proves that $\limsup_{x\to\infty} \varepsilon H + \tilde{H} < 0$ and Proposition 4.13 applies. □

Example 5 Consider the system (9) with $\alpha_i = 1$. Suppose that the conditions of Proposition 3.1 hold with

$$U(x) = 1 + \sum_i x_i^p$$

for some $p > 1$ and $\varphi = \eta = 0$. Assume also that

$$\mid F_i(x) \mid^q \leq cst\sqrt{U}$$

for some $q > 1$, and that F_i is bounded from below (i.e $\liminf_{x\to\infty} F_i(x) > -\infty$). Then the assumptions (30) and (31) of Theorem 5.1 hold true.

Example 6 (continuation of example 2) Using the notation of example 2, assume that

$$\mid F_i(x) \mid \leq cst(1+ \parallel x \parallel)$$

and

$$f_i(x_i) = r - bx_i$$

for some $b > 0$. Then the conditions (30) and (31) of Theorem 5.1 hold with

$$U(x) = 1 + \sum_i x_i^p$$

for all $p > 2$ and

$$\varphi(x) = \|x\|.$$

Call a point $x^* \in M$ *non-degenerate* if the matrix $a(x^*)$ is non-degenerate or, equivalently, if $\{\Sigma^1(x^*), \ldots, \Sigma^m(x^*)\}$ span $\mathbb{R}^n$.

Corollary 5.2 *Assume $I = \{1, \ldots, n\}$ (so that $M_0 = \partial\mathbb{R}^n_+$ and $M_+ = int(\mathbb{R}^n_+)$) and that the conditions (27), (28), (29) in Theorem 5.1 are satisfied. Assume furthermore that there exists $x^* \in \mathbf{\Gamma}_{M_+} \cap M_+$ which is non-degenerate. Then*

(i)

$$\mathcal{P}_{inv}(M_+) = \{\Pi\}$$

and for all $x \in M_+$

$$\lim_{t\to\infty} |P_t(x,\cdot) - \Pi|_{TV} = 0.$$

(ii) *Under the stronger conditions (30) and (31), there exist positive constant ε, θ, cst such that*

$$|P_t f(x) - \Pi f| \leq cst(1 + W_\theta(x))e^{-\lambda t}\|f\|_{W_\theta}$$

with

$$W_\theta(x) = \frac{U^\theta(x)}{(\prod_{i\in I} x_i^{p_i})^{\theta\varepsilon}}.$$

(iii) *If all the points in M_+ are non degenerate, then $\mathbf{\Gamma}_{M_+} \cap M_+ = M_+$.*

Proof: $(i), (ii)$. The non-degeneracy of $a(x^*)$ makes x^* a Doeblin point. Indeed, by Theorems 3.6 and 3.7 in Durrett [23], relying on Dynkin [24], there exist a open ball D centered at x^* and a positive function $q_t(x,y)$ continuous in $t > 0, x, y \in D$, such that if $f : \overline{D} \mapsto \mathbb{R}$ is continuous with $f|_{\partial D} = 0$,

$$\mathsf{E}_x(f(X_t)\mathbf{1}_{\tau>t}) = \int_D q_t(x,y)f(y)dy \tag{33}$$

for all $t > 0, x \in D$, where $\tau = \inf\{t \geq 0\ : X_t \notin D\}$. Hence, (23) holds with $\xi(dy) = \varepsilon \mathbf{1}_U(y)dy$, for some $\varepsilon > 0$ and $U \subset D$ a closed ball around x^*. The result then follows from Theorems 4.9 (respectively 4.12) and 5.1.

(iii). For all $x \in M_+, \mathbf{\Gamma}_x$ is a closed set containing x. If every point in M_+ is non-degenerate, the proof of (i) shows that $\mathbf{\Gamma}_x \cap M_+$ is open. Hence, by connectedness of $M_+, \mathbf{\Gamma}_x \cap M_+ = M_+$. □

Example 7 (Two dimensional systems) To illustrate the results above we consider here a simple model involving two species in interaction having the form

$$\begin{cases} dx_1 = x_1(F_1(x) + \sigma_1(x)dB_t^1) \\ dx_2 = x_2(F_2(x) + \sigma_2(x)dB_t^2) \end{cases} \tag{34}$$

where F_i, σ_i are smooth, σ_i is positive and bounded, $(B_t^1), (B_t^2)$ are two independent Brownian motions. We furthermore assume that the assumption of Proposition 3.1 as well as the conditions (27, 28) (or 31, 30) are satisfied (see for instance the examples 5 and 6).

We let $M_+ = \{x \in M \ : x_1 > 0, x_2 > 0\}$ and $M_0 = \{x \in M : \ x_1 x_2 = 0\}$.

We let
$$\lambda_i(x) = F_i(x) - \frac{\sigma_i^2(x)}{2}.$$
On the invariant face $x_2 = 0$ the process admits one ergodic probability given by the dirac at the origin $\delta_{0,0}$ and an invariant measure (non necessarily a probability) $h_1(x_1)dx_1\delta_0(dx_2)$ where

$$h_1(x_1) = \frac{2}{x_1^2\sigma_1^2(x_1,0)} \exp\{\int_r^{x_1} \frac{2F_1(u,0)}{u\sigma_1(u,0)^2} du\}\mathbf{1}_{x_1>0}$$

and $r > 0$ is an arbitrary number. This invariant measure is finite if only if the integrability condition

$$\frac{2F_1(0,0)}{\sigma_1^2(0,0)} > 1 \Leftrightarrow \lambda_1(0,0) > 0$$

holds true. Observe that this condition is exactly the persistence condition (5.1) of the process restricted to the face $\{x \in M : \ x_2 = 0\}$ with invariant set $\{0,0\}$. In this later case we let μ_1 denote the ergodic probability obtained by normalizing h_1. That is

$$\mu_1(dx_1dx_2) = \frac{h_1(x_1)}{\int_0^\infty h_1(u)du} dx_1\delta_0(dx_2).$$

We define μ_2 similarly. In summary,

$$\mathsf{P}_{erg}(M_0) = \begin{cases} \{\delta_{0,0}\} \text{ if } \lambda_1(0,0) < 0 \text{ and } \lambda_2(0,0) < 0, \\ \{\delta_{0,0}, \mu_1\} \text{ if } \lambda_1(0,0) > 0 \text{ and } \lambda_2(0,0) < 0, \\ \{\delta_{0,0}, \mu_2\} \text{ if } \lambda_1(0,0) < 0 \text{ and } \lambda_2(0,0) > 0, \\ \{\delta_{0,0}, \mu_1, \mu_2\} \text{ if } \lambda_1(0,0) > 0 \text{ and } \lambda_2(0,0) > 0. \end{cases}$$

Therefore, the persistence condition (29) is satisfied in one of the three following cases:

(i) $\lambda_1(0,0) > 0, \lambda_2(0,0) < 0$, and $\mu_1(\lambda_2) > 0$, or

(ii) $\lambda_1(0,0) < 0, \lambda_2(0,0) > 0$, and $\mu_2(\lambda_1) > 0$, or

(iii) $\lambda_1(0,0) > 0, \lambda_2(0,0) > 0, \mu_1(\lambda_2) > 0$, and $\mu_2(\lambda_1) > 0$.

In each case the conclusions of Corollary 5.2 hold. (Compare to section 4.3 of [70] and to Section 2.1 of [37]).

Example 8 (Example 7 continued, Randomness promotes persistence) Suppose that the noise term σ_i in equation (34) takes the form

$$\sigma_i = \varepsilon s_i$$

for some $0 < \varepsilon << 1$ with $s_i > 0$ and bounded.

Let $V_1, V_2 :]0,\infty[\mapsto \mathbb{R}$ be the maps defined by

$$V_1(t) = \int_r^t \frac{F_1(u,0)}{us_1^2(u,0)} du \text{ and } V_2(t) = \int_r^t \frac{F_2(0,u)}{us_2^2(0,u)} du.$$

If $F_i(0,0) > 0$, V_i achieves its maximum at a point $x_i^* > 0$. Assume for simplicity that such a maximum is unique. Note that $F_1(x_1^*,0) = 0$ (similarly $F_2(0,x_2^*) = 0$) and that $\frac{\partial F_1}{\partial x_1}(x_1^*,0) \leq 0$ (similarly $\frac{\partial F_2}{\partial x_2}(0,x_2^*) \leq 0$).

If $F_1(0,0) > 0$ (respectively $F_2(0,0) > 0$) the probability μ_1 (respectively μ_2) converges when $\varepsilon \to 0$ (in the weak * sense) toward the dirac measure at $(x_1^*,0)$ (respectively $(0,x_2^*)$.) Therefore, using the results described in example 7, one see that the process is stochastically persistent for every $\varepsilon > 0$ sufficiently small, provided one of the following conditions hold:

(i) $F_1(0,0) > 0, F_2(0,0) < 0$ and $F_2(x_1^*,0) > 0$, or

(ii) $F_1(0,0) < 0, F_2(0,0) > 0$ and $F_1(0,x_2^*) > 0$, or

(iii) $F_1(0,0) > 0, F_2(0,0) > 0, F_2(x_1^*,0) > 0$ and $F_1(0,x_2^*) > 0$

An interesting consequence of this result is that an arbitrary small random perturbation of a non persistent deterministic ecological ODE can be stochastically persistent. Indeed condition (i) above simply means that the origin is a saddle point (for the ode obtained with $\varepsilon = 0$) which stable (respectively unstable) manifold is the axis $x_1 = 0$ (respectively $x_2 = 0$) and the point $(x_1^*,0)$ a saddle point which stable manifold is the axis $x_2 = 0$. However there may exist other equilibria on the boundary including sinks or saddle points.

5.1 Degenerate ecological SDEs

We discuss here the situation where (9) is a degenerate SDE. This is motivated by models for which the noise only affects certain variables. We assume throughout the section that the vector fields F and $\Sigma^j, j = 1, \ldots, m$ are C^∞.

Rewrite the stochastic differential equation (9) using the Stratonovich formalism as

$$dx_t = S^0(x_t)dt + \sum_{j=1}^{m} S^j(x_t) \circ dB_t^j \tag{35}$$

where for all $j = 1 \ldots m$,

$$S_i^j(x) = x_i^\alpha \Sigma_i^j(x), i = 1 \ldots n$$

and

$$S_i^0(x) = x_i^{\alpha_i} F_i(x) - \frac{1}{2} \sum_{j=1}^{m} \sum_{k=1}^{n} \frac{\partial S_i^j}{\partial x_k}(x) S_k^j(x), i = 1 \ldots n.$$

Associated to this system is the deterministic control system

$$\dot{y}(t) = S^0(y(t)) + \sum_{j=1}^{m} u^j(t) S^j(y(t)) \tag{36}$$

where the control function $u = (u^1, \ldots, u^m) : \mathbb{R}_+ \mapsto \mathbb{R}^m$, can be chosen to be piecewise continuous. Given such a control function, we let $y(u, x, \cdot)$ denote the maximal solution[4] to (36) starting from x (i.e $y(u, x, 0) = x$). The following proposition easily follows from the the celebrated Strook and Varadhan's support theorem [74] (see also Theorem 8.1, Chapter VI in [47]. Recall that we let $\mathbf{\Gamma}_x$ denote the accessible set from x (as defined in section 4.2).

Proposition 5.3 *Let $x \in M$. Point $p \in M$ lies in $\mathbf{\Gamma}_x$ if and only if for every neighborhood O of p there exist a control u such that $y(u, x, \cdot)$ meets O (i.e $y(u, x, t) \in O$ for some $t \geq 0$).*

Proof: If the vector fields S^j were bounded with bounded (first and second) derivatives, this would follow directly from the support theorem (see Theorem 8.1, Chapter VI in [47]). To handle the fact that the S^j are typically unbounded we use a localization argument relying on the existence

[4]Note that there is no assumption here that the vector fields S^j are globally integrable.

of the Lyapunov function U assumed in proposition 3.1. Let $S^{j,n}$ be a smooth vector field with compact support coinciding with S^j on the set $U_n = \{x \in M \ : U(x) < n\}$. Let $\mathbb{P}_x^n$ be for the law of the process starting from x solution to the SDE obtained by replacing S^j by $S^{j,n}$ in (35). Let $y^n(x,u,\cdot)$ be defined like $y(x,u,\cdot)$ when S^j is replaced by $S^{j,n}$ in (36). Let $\tau_n = \inf\{t \geq 0 \ : \ U(X_t) \geq n\}$. The assumptions on U imply that $\lim_{n\to\infty} \mathbb{P}_x(\tau_n > t) = 1$ (see the proof of Proposition 3.1, equation (64)). Thus, for every open set $O \subset M$,

$$\mathbb{P}_x(X_t \in O) > 0 \Leftrightarrow \exists n \ \mathbb{P}_x(X_t \in O; \tau_n > t) > 0 \Leftrightarrow \exists n \ \mathbb{P}_x^n(X_t \in O; \tau_n > t) > 0.$$

By the support theorem,

$$\begin{aligned} \mathbb{P}_x^n(X_t \in O; \tau_n > t) > 0 &\Leftrightarrow \exists u \, y^n(u,x,[0,t]) \subset U_n \text{ and } y^n(u,x,t) \in O \\ &\Leftrightarrow \exists u \, y(u,x,[0,t]) \subset U_n \text{ and } y(u,x,t) \in O \end{aligned}$$

because $S^j = S^{j,n}$ on U_n. Therefore

$$\mathbb{P}_x(X_t \in O) > 0 \Leftrightarrow \exists u \, y(u,x,t) \in O.$$

This proves the result. $\square$

The local *ellipticity* condition given by the non degeneracy of $a(x^*)$ in Corollary 5.2 can be weakened and replaced by a local *hypoellipticity* condition.

Recall that the Lie bracket of two smooth vector fields $Y, Z : \mathbb{R}^n \mapsto \mathbb{R}^n$ is the vector field defined as

$$[Y,Z](x) = DZ(x)Y(x) - DY(x)Z(x).$$

Given a family $\mathcal{X}$ of smooth vector fields on $\mathbb{R}^n$, we let $[\mathcal{X}]_k, k \in \mathbb{N}$, and $[\mathcal{X}]$ denote the set of vector fields defined by $[\mathcal{X}]_0 = \mathcal{X}$,

$$[\mathcal{X}]_{k+1} = [\mathcal{X}]_k \cup \{[Y,Z] \ : Y, Z \in [\mathcal{X}]_k\}$$

and $[\mathcal{X}] = \cup_k [\mathcal{X}]_k$. We also let $[\mathcal{X}](x) = \{Y(x) \ : Y \in [\mathcal{X}]\}$.

Consider again the SDE (9) (or equivalently 35). We say that $x^* \in M$ satisfies the *Hörmander condition* (respectively the *strong Hörmander condition*) if $[\{S^0, \ldots, S^m\}](x^*)$ (respectively

$$\{S^1(x^*), \ldots, S^m(x^*)\} \cup \{[Y,Z](x^*) : \ Y, Z \in [\{S^0, \ldots, S^m\}]\})$$

spans $\mathbb{R}^n$.

The next corollary just states that the local ellipticity condition in Corollary 5.2 can be weakened to a local hypoellipticity.

Corollary 5.4 *Assume $I = \{1, \dots, n\}$, so that $M_0 = \partial\mathbb{R}^n_+$ and $M_+ = int(\mathbb{R}^n_+)$, and that the conditions (27, 28) and (29) of Theorem 5.1 are satisfied.*

(i) *If there exists $x^* \in \mathbf{\Gamma}_{M_+} \cap M_+$ which satisfies the Hörmander condition, then*

$$\mathcal{P}_{inv}(M_+) = \{\Pi\},$$

where $\Pi << \lambda$ (the Lebesgue measure on $\mathbb{R}^n$), and for all $x \in M_+$

$$\lim_{t\to\infty} \Pi^x_t = \Pi$$

P *a.s*

(ii) *If the Hörmander condition at x^* is strengthened to the strong Hörmander condition, then (P_t) converge to Π in total variation (like in Corollary 5.2 (i)). Under the stronger conditions (30) and (31), the convergence is exponential ((like in Corollary 5.2 (ii)).*

(iii) *If for all $x \in M_+$ $[\{S^1, \dots S^m\}](x)$ spans $\mathbb{R}^n$, then $\mathbf{\Gamma}_{M_+} \cap M_+ = M_+$.*

Proof: (i) Let D be a domain (connected open set) containing x^*, relatively compact, and small enough so that $[\{S^0, \dots, S^m\}](x)$ spans $\mathbb{R}^n$ for each $x \in \bar{D}$. First assume that

(a) For each $x \in \bar{D}$ $\sum_{i=1}^m \|S^i(x)\| \neq 0$;

(b) For each $x \in \partial D = \bar{D} \setminus D$ there exists a vector u normal to $\bar{D}$ such that $\sum_{i=1}^m \langle S^i(x), u\rangle^2 > 0$.

Under these assumptions, by a Theorem of Bony ([13], Theorem 6.1), there exists a kernel $G : \bar{D} \times \bar{D} \mapsto \mathbb{R}_+$, smooth on $D \times D \setminus \{(x, x) \,: x \in D\}$ such that: For each $f \in C_b(\bar{D})$, there exists a unique $g \in C_b(\bar{D})$ solution to the Dirichlet problem

$$\begin{cases} Lg - g = -f \text{ on } D(\text{ in the sense of distributions}) \\ g|_{\partial D} = 0; \end{cases}$$

and $g(x) = Gf(x) := \int G(x, y)f(y)dy$. Furthermore, if f is smooth on D so is g.

Note that, by continuity of G off the diagonal, there exist disjoint open sets $U, V \subset D$, with $x^* \in U$ and $\delta > 0$ such that $G(x, y) \geq \delta$ on $U \times V$.

Let $\tau = \inf\{t > 0\, X_t \notin D\}$. For f smooth on D, Ito's formula shows that,

$$(e^{-t\wedge\tau} g(X_{t\wedge\tau}) + \int_0^{t\wedge\tau} e^{-s} f(X_s) ds)$$

is a local martingale. Being bounded, it is a uniformly integrable martingale. Thus,

$$\mathsf{E}_x(\int_0^\tau e^{-s} f(X_s) ds) = Gf(x).$$

Let $R(x,\cdot) = \int_0^\infty e^{-t} P_t(x,\cdot)$. It follows that for all $x \in U$

$$R(x,dy) \geq \delta \mathbf{1}_V(dy)$$

and the result follows from Proposition 4.7.

It remains to explain how we can choose D to ensure that conditions (a) and (b) above are satisfied. We assume here that $n \geq 2$. For $n = 1$ the proof is left to the reader. If $\sum_{i\geq 1} \|S^i(x^*)\| > 0$, then (a) holds provided D is small enough. If $\sum_{i\geq 1} \|S^i(x^*)\| = 0$, set $y^* = \Phi_t^0(x^*)$ where $\{\Phi_t^0\}$ is the local flow induced by S^0. We claim that, for $t > 0$ small enough, $y^* \in D$ and $\sum_{i\geq 1} \|S^i(y^*)\| > 0$. Since y^* is accessible, it then suffices to replace x^* by y^* and D by a neighborhood of y^*. To prove this claim, assume to the contrary, that $S^i(\Phi_t^0(x^*)) = 0$ for all $0 < t < \varepsilon$ and $i = 1,\ldots,m$. Then

$$0 = DS^i(\Phi_t^0(x^*))\frac{d}{dt}\Phi_t^0(x^*) = DS^i(y^*)S^0(y^*) = [S^0, S^i](y^*).$$

Similarly $Z(y^*) = 0$ for all $Z \in \{[S^0,\ldots,S^m]\} \setminus \{S^0\}$. A contradiction.

For condition (b), we can assume (by condition (a)) that $S^1(x^*) \neq 0$ and without loss of generality that $\frac{S^1(x^*)}{\|S^1(x^*)\|} = e_1$ the first vector in the canonical basis of $\mathbb{R}^n$. Let, for $\varepsilon > 0$, small enough $D_\varepsilon = \{x \in \mathbb{R}^n\, \|x - x^*\|_1 < \varepsilon\}$ where $\|u\|_1 = \sum_{i=1}^n |u_i|$. For $x \in \partial D_\varepsilon$ let u_x be the vector defined by $u_{x,i} = \frac{x_i - x_i^*}{|x_i - x_i^*|}$ if $x_i \neq x_i^*$ and $u_i^* = 1$ otherwise. Vector u_x is normal to ∂D and $\langle e_1, u_x\rangle^2 = 1$. Hence, for ε small enough $\langle S^1(x), u_x\rangle^2 > 0$ for all $x \in \partial D$. It suffices to replace D by D_ε and (b) is satisfied.

(ii) Under the strong Hörmander condition, the law of (X_t) killed at D (see Ichihara and Kunita [46]) has a density $q_t(x,y)$ which is C^∞ in $t > 0, x, y \in D$. Choose $y^* \in D$ and $t^* > 0$ such that $q_{t^*}(x^*, y^*) > 0$ (such a (t^*, y^*) exists for otherwise τ would be almost surely 0 contradicting the continuity of paths). The end of the proof is then identical to the proof of Corollary 5.2.

(iii) follows from Chow's Theorem. $\square$

Remark 16 In case all the points in M_+ satisfy the Hörmander condition, the invariant measure Π in the previous corollary has a C^∞ density by hypoellipticity of L^* (the formal adjoint of L).

5.2 A stochastic Rosenzweig-MacArthur model

As an illustration of the previous result, we consider here the Rosenzweig-MacArthur model described in Example 3, under the assumption that only the prey-variable is subjected to some small environmental fluctuation. That is

$$\begin{cases} dx_1 = x_1(F_1(x_1, x_2)dt + \varepsilon dB_t) \\ dx_2 = x_2 F_2(x_1, x_2)dt \end{cases} \tag{37}$$

where

$$F_1(x) = 1 - \frac{x_1}{\kappa} - \frac{x_2}{1 + x_1}, F_2(x) = -\alpha + \frac{x_1}{1 + x_1},\ \alpha, \kappa > 0.$$

We let $M_0 = \partial\mathbb{R}^2_+, M_+ = Int(\mathbb{R}^2_+)$.

For $0 < \varepsilon^2 < 2$, let $k = \frac{2}{\varepsilon^2} - 1, \theta = \frac{\varepsilon^2\kappa}{2}$ and let

$$\gamma_{\varepsilon,\kappa}(x) = \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\Gamma(k)\theta^k}$$

be the density of a Γ distribution with parameters k, θ. Set

$$\Lambda(\varepsilon, \alpha, \kappa) = \int_0^\infty \frac{x}{1 + x}\gamma_{\varepsilon,\kappa}(x)dx - \alpha.$$

A rough estimate of $\Lambda(\varepsilon, \alpha, \kappa)$ is

$$\frac{\kappa(1 - \frac{\varepsilon^2}{2})}{1 + \kappa(1 - \frac{\varepsilon^2}{2})} - \kappa\frac{\varepsilon}{4}\sqrt{1 - \frac{\varepsilon^2}{2}} - \alpha \leq \Lambda(\varepsilon, \alpha, \kappa) \leq \frac{\kappa(1 - \frac{\varepsilon^2}{2})}{1 + \kappa(1 - \frac{\varepsilon^2}{2})} - \alpha.$$

The right hand side inequality follows from Jensen inequality and the fact that $\gamma_{\varepsilon,\kappa}$ has mean $k\theta = \kappa(1 - \frac{\varepsilon^2}{2})$. The left hand side follows from the fact that $x \mapsto \frac{x}{1+x}$ is $1-$ Lipschitz, Cauchy Schwarz inequality and the fact that $\gamma_{\varepsilon,\kappa}$ has variance $k\theta^2 = \kappa^2\frac{\varepsilon^2}{2}(1 - \frac{\varepsilon^2}{2})$.

Theorem 5.5 *System (37) behaves as follows:*

(i) *If $\Lambda(\varepsilon, \alpha, \kappa) > 0$ (in particular $0 < \varepsilon^2 < 2$), then for all $x \in M^+$ (Π_t^x) (respectively $(P_t(x, \cdot))$ converges almost surely (respectively in total variation) toward a unique measure Π (depending on ε). Furthermore, Π has a smooth density (with respect to Lebesgue measure) strictly positive over M_+.*

(ii) *If $\Lambda(\varepsilon, \alpha, \kappa) < 0$ (in particular $0 < \varepsilon^2 < 2$), then $x_2(t) \to 0$ a.s and for all $x \in M^+$ $x(t) \Rightarrow \gamma_{\varepsilon,\kappa}(dx_1) \otimes \delta_0(dx_2)$.*

(iii) *If $\varepsilon^2 > 2$, then $x(t) \to 0$ almost surely.*

Proof: We only prove (i). Assertion (ii) and (iii) will be proved in Part II. Fix $n > 2$. We first notice that the assumptions of Proposition 3.1 are satisfied with $U(x) = (x_1 + x_2)^n, \eta = 0$ and $\varphi = 0$. Indeed,

$$LU(x) = n[(x_1 + x_2)^{n-1}(x_1(1 - \frac{x_1}{\kappa}) - \alpha x_2) + \frac{n-1}{2}(x_1 + x_2)^{n-2}\varepsilon^2 x_1^2]$$

$$\leq n(x_1 + x_2)^{n-1}[x_1(1 + \frac{n-1}{2}\varepsilon^2 - \frac{x_1}{\kappa}) - \alpha x_2]$$

$$\leq n(x_1 + x_2)^{n-1}(-\alpha(x_1 + x_2) + \beta)$$

for some $\beta > 0$. Thus

$$LU \leq -\alpha U + \tilde{\beta}$$

for some $\tilde{\beta} > 0$. Also

$$\Gamma_L(U)(x) = (n(x_1 + x_2)^{n-1}\varepsilon x_1)^2 \leq n^2\varepsilon^2 U^2(x).$$

With such a function U the conditions (27) and (28) of Theorem 5.1 are clearly satisfied.

Reasoning like in Example 7 we see that for $\varepsilon^2 < 2$, $\mathsf{P}_{erg}(M_0) = \{\delta_{0,0}, \mu_1\}$ where $\mu_1(dx_1 dx_2) = \gamma_{\varepsilon,\kappa}(x_1)dx_1\delta_0(dx_2)$ and that the persistence condition (29) is given by $\Lambda(\varepsilon, \alpha, \kappa) > 0$.

The vector fields S^0, S^1 (in equation (35)) write

$$S^0(x_1, x_2) = \begin{pmatrix} x_1(F_1(x_1, x_2) - \varepsilon^2/2)) \\ x_2 F_2(x_1, x_2) \end{pmatrix}; \; S^1(x_1, x_2) = \begin{pmatrix} x_1\varepsilon \\ 0 \end{pmatrix}.$$

Thus, by a simple computation,

$$det([S^0, S^1](x), S^1(x)) = \varepsilon^2 x_1^2 x_2 \frac{\partial F_2}{\partial x_1} = \frac{\varepsilon^2 x_1^2 x_2}{(1+x_1)^2}.$$

This shows that the strong Hörmander condition holds at every point $x \in M_+$. To prove the claim, it remains to show that $\mathbf{\Gamma}_{M_+}$ contains M_+. Existence and convergence to Π will then follow from Corollary 5.4 (ii). Smoothness of the density from Hypoellipticity of the adjoint L^* (see e.g the reasoning in [46] before the proof of Proposition 5.1) and positivity of the density from the fact that $\mathbf{\Gamma}_{M+}$ is the support of Π.

Introduce the new control variable $v = \varepsilon u - \varepsilon^2/2$. Then, the control system (36) rewrites

$$\dot{x_1} = x_1(F_1(x) + v), \dot{x_2} = x_2 F_2(x).$$

Let L be the line $x_1 = \frac{\alpha}{1-\alpha}$. That is $F_2(x) = 0$. Let (P_v) be the parabola $(1+v-\frac{x_1}{\kappa})(1+x_1) = x_2$. That if $F_1(x)+v = 0$. Let $z = (z_1, z_2) \in M_+$ and O_z a neighborhood of z. Choose v^* large enough so that z is below (P_{v^*}) and the point at which (P_{v^*}) reaches its maximum is $> \frac{\alpha}{1-\alpha}$. It is not hard to verify that there exists a neighborhood O of the origin such that for all $x \in O \cap M_+$ $t \mapsto y(v^*, x, t)$ crosses (P_{v^*}) near $(v^*, 0)$ then remains above (P_{v^*}) and then crosses (L). In particular, it crosses the line $x_2 = z_2$. Given any $x \in M_+$ use a piecewise constant control $v(t)$ as follows: $v(t) = -1$ until $t \mapsto y(v, x, t)$ enters O. Then $v(t) = v^*$ until $t \mapsto y(v, x, t)$ crosses the horizontal line $x_2 = z_2$. Then $v(t) = -R$ where R is large enough so that $t \mapsto y(v, x, t)$ eventually enters O_z. By Proposition 5.3 this proves that $z \in \mathbf{\Gamma}_x$. $\square$

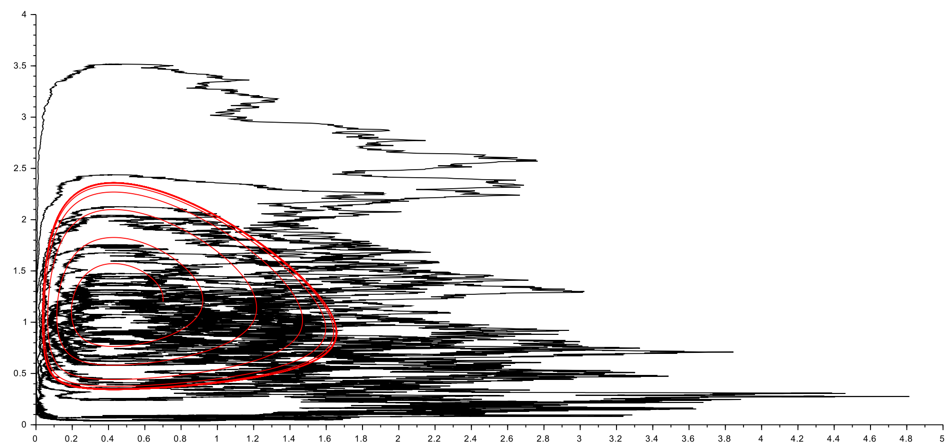

Figure 1: $\alpha = 0,3; \kappa = 2,5 : \varepsilon = 0,6$

Figure 1 obtained in Scilab by Edouard Strickler, illustrates the behavior of the process when $\Lambda(\varepsilon, \alpha, \kappa) > 0$. The red trajectory is a trajectory of the unperturbed system.

6 Application to Random Ecological ODEs (ii)

Here we consider the random ecological ODEs introduced in section 3.3 and use the same notation. Recall that the state space has the form $M = B \times \{1, \ldots, m\}$, with B compact, and that for $I \subset \{1, \ldots, n\}$ we let

$$M_0^I = \{(x, u) \in M \ : \prod_{i \in I} x_i = 0\}.$$

The *invasion rate* of species i with respect to (x, u) is defined as $\lambda_i(x, u) = F_i^j(x)$ and invasion rate of species i with respect to $\mu \in \mathcal{P}_{erg}(M_0^I)$ as

$$\lambda_i(\mu) = \sum_j \int_B F_i^j(x) d\mu^j(x) \tag{38}$$

where $\mu^j(A) := \mu(A \times \{j\})$.

Theorem 6.1 *Assume that that there exist positive number $\{p_i\}_{i\in I}$ such that for all $\mu \in \mathcal{P}_{erg}(M_0^I)$*

$$\sum_{i\in I} p_i\lambda_i(\mu) > 0.$$

Then the process given by (20) is H persistent with respect to M_0^I.

Proof: The proof is similar to the proof of Theorem 5.1. If $(x,u) \in M_+$ (respectively $(x,u) \in M$) let $V(x)$ (respectively $V_n(x)$) be defined exactly as in the proof of Theorem 5.1. See V and V_n as functions of (x,u) (i.e set $V(x,u) := V(x), V_n(x,u) := V_n(x)$) and let $H(x,u) = -\sum_{i\in I} p_i\lambda_i(x,u)$. Then $V_n \in \mathcal{D}^2(\mathcal{L})$, $V_n = V$ and $\mathcal{L}(V_n) = H$ on the set $K_n = \{(x,u) \in M_+ : V(x,u) \leq n, \sum_i x_i \leq n\}$. Furthermore $\Gamma(V_n)(x,u) = 0$ (because $V_n(x,u)$ doesn't depend on u) so that assumption (i) of definition 4 is satisfied □

Associated to (20) is the control system

$$\dot{y}(t) = \sum_{j=1}^{m} u^j(t)G^j(y(t)) \tag{39}$$

where the control function $u = (u^1,\ldots,u^m) : \mathbb{R}_+ \mapsto \mathbb{R}^n$ can be chosen to be piecewise continuous with values either in $\{e_1,\ldots,e_n\}$, the canonical basis of $\mathbb{R}^n$; or in Δ^{n-1}, the unit simplex of $\mathbb{R}^n$. The solution to (39) starting from x is denoted $t \mapsto y(u,x,t)$.

The following proposition is analogous to Proposition 5.3 and follows from the support theorem established in ([9], Theorem 3.4). Note that this support theorem is phrased in terms of the differential inclusion whose set valued vector field is given by the convex hull of $\{G^1,\ldots,G^m\}$ but the link with the control system (39) is spelled out, for instance, in ([5], Theorem 2.2).

Proposition 6.2 *Let $(x,i) \in M$. Point $(p,j) \in M$ lies in $\mathbf{\Gamma}_{(x,i)}$ if and only if for every neighborhood O of p there exists a control u such that $y(u,x,\cdot)$ meets O (i.e $y(u,x,t) \in O$ for some $t \geq 0$).*

By analogy with the terminology used for SDE's in section 5.1, we say that $(x,i) \in M$ satisfies the *Hörmander* or *weak bracket* (the terminology coined in [9]) condition, respectively the *strong Hörmander* or *strong bracket condition* if $[\{G^1,\ldots,G^m\}](x)$, respectively

$$\{G^i(x) - G^j(x) \ : i,j = 1,\ldots,m\} \cup \{[Y,Z](x) : \ Y,Z \in [\{G^1,\ldots,G^m\}]\}$$

spans $\mathbb{R}^n$. These two conditions are named A (for the stronger) and B (for the weaker) in [3].

Corollary 6.3 *Assume* $I = \{1, \ldots, n\}$, *and that the condition of Theorem 6.1 holds. Assume that there exists* $(x^*, j) \in \mathbf{\Gamma}_{M_+} \cap M_+$ *which satisfies the Hörmander condition. Then*

(i)

$$\mathcal{P}_{inv}(M_+) = \{\Pi\},$$

where $\Pi << \lambda$ *(the Lebesgue measure on* M*), and for all* $(x, j) \in M_+$

$$\lim_{t\to\infty} \Pi_t^{(x,j)} = \Pi$$

P *a.s*

(ii) *Assume in addition that, either*

(a) *the Hörmander condition at* (x^*, j) *is strengthened to the strong Hörmander condition, or*

(b) *There exist* $\alpha_j, \ldots, \alpha_m \in \mathbb{R}$ *with* $\sum_j \alpha_j = 1$ *and* $e \in \mathbf{\Gamma}_{M^+} \cap M^+$ *for which*

$$\sum_{j=1}^{m} \alpha_j G^j(e) = 0.$$

Then for all $(x, j) \in M_+$

$$\lim_{t\to\infty} |P_t((x, j), \cdot) - \Pi|_{TV} \leq cst(1 + W_\theta(x))e^{-\lambda t}.$$

for some $\theta, \lambda > 0$. *Here*

$$W_\theta(x) = e^{\theta[\max(\sum_i p_i h_i(x_i), 1)]}$$

where $h_i(u) = -\log(u)$ *for* $\alpha_i = 1$ *and* $h_i(u) = \frac{u^{1-\alpha_i}}{1-\alpha_i}$ *if* $\alpha_i > 1$.

Proof: Under condition (i) $\mathcal{P}_{inv}(M_+)$ has at most cardinal one, as proved in [3] for constant rates a_{ij} and in [9] for nonconstant rates $(a_{ij}(x))$. Note that for constant rates [3] actually prove that the condition of Proposition 4.7 are satisfied with $\gamma(dt) = e^{-t}dt$. The result then follows from Corollary 4.5. Under condition $(ii)a)$, it follows again from [3] (for constant rates) and [9] for nonconstant rates that x^* is a Doeblin point. Under condition $(ii), b)$ this follows from a result recently proved in [7] strongly inspired by the work of [52]. The result then follows from Theorem 4.10. □

Remark 17 In contrast with the SDE situation (see remark 16), the question of the smoothness of the density of the invariant measure remains largely an open problem (see [2], [1] for some results in dimension one and two).

6.1 A May-Leonard System with random switching

The goal of this section is twofold: Illustrate the preceding results and provides a simple example, albeit non trivial, for which the H-persistence machinery applies to a situation where the extinction set is not just the boundary of $\mathbb{R}^n_+$.

Let $r : \mathbb{R}^3_+ \mapsto \mathbb{R}^*_+$ be a smooth function, (α, β) a pair or parameters - called an *environment* -satisfying

$$0 < \beta < 1 < \alpha, \tag{40}$$

and let G be the vector field on $\mathbb{R}^3_+$ defined as $G_i(x) = x_i F_i(x)$, with

$$F(x) = r(x) \begin{cases} (1 - x_1 - \alpha x_2 - \beta x_3) \\ (1 - \beta x_1 - x_2 - \alpha x_3) \\ (1 - \alpha x_1 - \beta x_2 - x_3) \end{cases} \tag{41}$$

When $r := 1$ we recover the celebrated model introduced by May and Leonard [56] in 1975. A nonconstant r has no effect on the phase portrait of G (it only changes the velocity) but will have some on the persistence properties of the process obtained by random switching of the parameters.

Before considering such a process, we first recall some basic properties of the dynamics induced by G.

Background We let $\Phi = \{\Phi_t\}$ denote the local solution flow in $\mathbb{R}^3_+$ to the differential equation $\dot{x} = G(x)$. Here, the terminology *trajectories, equilibria, limit points* etc. refer to trajectories, equilibria, limit points of Φ. Throughout we let

$$N(x) = x_1 + x_2 + x_3$$

and

$$\Delta = \{x \in \mathbb{R}^3_+ : N(x) = 1\}$$

denote the unit simplex.

Vector field G has 5 equilibria, the origin 0 which is a source, the canonical basis vectors e_1, e_2, e_3 which are saddle points, and the interior equilibrium

$$x_* = \frac{e_1 + e_2 + e_3}{\alpha + \beta + 1}.$$

The diagonal

$$D = \{x \in \mathbb{R}^3_+ \ : x_1 = x_2 = x_3\}$$

is invariant and on D every nonzero trajectory converges to x_*.

On the face $\mathbb{R}_+ \times \mathbb{R}^*_+ \times \{0\}$ every trajectory converges to e_2. In words, *species* 2 *beats species* 1 *in absence of species* 3. Similarly, species 1 beats 2 in absence of 3 and 3 beats 2 in absence of 1. This makes the set

$$\Upsilon = W^s(e_1) \cup W^s(e_2) \cup W^s(e_3)$$

an heteroclinic cycle; where $W^s(e_i)$ stands for the stable manifold of e_i.

The global dynamics of G can now be described :

- If $\alpha + \beta < 2$, x_* is a sink and every trajectory starting from $\mathbb{R}^3_+ \setminus \partial\mathbb{R}^3_+$ converges to x_*.

- If $\alpha + \beta > 2$, x_* is a saddle whose stable manifold is $D \setminus \{0\}$ and every trajectory starting from $\mathbb{R}^3_+ \setminus (\partial\mathbb{R}^3_+ \cup D)$ has Υ as omega limit cycle.

- If $\alpha + \beta = 2$, Δ is invariant and attracts every nonzero trajectory. In this case $\Upsilon = \partial\Delta$ and on $\Delta \setminus (\{x_*\} \cup \Upsilon)$ every trajectory is periodic.

All these properties are proved in Section 5.5 of [42].

When $\alpha+\beta \neq 2$, Δ is no longer invariant but general results on competitive systems first developed by Hirsch (see in particular [40], Theorem 1.7 or Hirsch and Smith [39], Theorem 3.18) imply that

- There exists a compact invariant set Σ, unordered, homeomorphic to Δ by radial projection $x \mapsto \frac{x}{N(x)}$, such that $\Phi_t(x) \to \Sigma$ as $t \to \infty$, for all $x \neq 0$. Also, $\Sigma \cap \partial\mathbb{R}^3_+ = \Upsilon$ and $\Sigma \cap (\mathbb{R}^3_+ \setminus \partial\mathbb{R}^3_+)$ is a Lipschitz manifold.

Unordered means that if $x, y \in \Sigma$ and $y - x \in \mathbb{R}^3_+$, then $x = y$. The set Σ is called the *carrying simplex*, a term coined by M. Zeeman [77], and can be characterized as the boundary (in $\mathbb{R}^3_+$) of the basin of repulsion of ∞ or equivalently, the boundary of the basin of repulsion of the origin. That is

$$\Sigma = \partial_{\mathbb{R}^3_+} R(\infty) = \partial_{\mathbb{R}^3_+} R(0)$$

where

$$R(\infty) = \{x \in \mathbb{R}^3_+ : \lim_{t \to -\infty} \|\Phi_t(x)\| = \infty\} \text{ and } R(0) = \{x \in \mathbb{R}^3_+ : \lim_{t \to -\infty} \Phi_t(x) = 0\}.$$

Smoothness properties of Σ have been investigated in several papers (see in particular Mierczynski [58, 59]). Further properties of the carrying simplex for Lotka Volterra systems are discussed in Zeeman [77].

Clearly

$$\dot{N} = r(x)(N - (x_1^2 + x_2^2 + x_3^2) - (\alpha + \beta)(x_1x_2 + x_1x_3 + x_2x_3)) \tag{42}$$

from which it follows that for $\alpha + \beta < 2$ (respectively $\alpha + \beta > 2$) $\dot{N} > N(1-N)$(respectively $\dot{N} < N(1-N)$) whenever $x_1x_2 + x_2x_3 + x_1x_3 \neq 0$. As a consequence,

- If $\alpha + \beta < 2$, Σ is *above* Δ. That is $N(x) > 1$ for all $x \in \Sigma \setminus \{e_1, e_2, e_3\}$;
- If $\alpha + \beta > 2$, Σ is *below* Δ. That is $N(x) < 1$ for all $x \in \Sigma \setminus \{e_1, e_2, e_3\}$.

Random switching Let (α_1, β_1) and (α_2, β_2) be two environments - as defined by equation (40) - such that

$$\alpha_1 + \beta_1 > 2 \text{ and } \alpha_2 + \beta_2 < 2.$$

For each j we let G^j denote the vector field defined like G in environment (α^j, β^j), and $\Phi^j, x_*^j, \Upsilon^j, \Sigma^j, etc.$ the corresponding flow, interior equilibrium, heteroclinic cycle, carrying simplex, etc.

In view of the preceding discussion, for each $x \in \Delta$ the line $\mathbb{R}^+ x$ meets Σ^j in a single point $\varsigma_j(x)$. If $x \notin \{e_1, e_2, e_3\}$ $N(\varsigma_1(x)) < 1 < N(\varsigma_2(x))$ while for $x \in \{e_1, e_2, e_3\}, \varsigma_i(x) = x$. Define the *cell bordered by* Σ^1, Σ^2 as

$$\mathbf{C}(\Sigma^1, \Sigma^2) = \{t\varsigma_1(x) + (1-t)\varsigma_2(x) \ : 0 \leq t \leq 1, x \in \Delta\}.$$

This set is homeomorphic to the closed unit ball in $\mathbb{R}^3$ and its boundary (in $\mathbb{R}^3_+$) is the union of the carrying simplices Σ^1 and Σ^2. It will characterize the support of the persistent measure (when there is such a measure).

Fix $0 < \eta < \frac{\alpha_2 + \beta_2}{6}$ and set

$$B = \{x \in \mathbb{R}^3_+ \ : 3\eta \leq N(x) \leq 3\}.$$

By equation (42), B is positively invariant by G^1 and G^2.

Consider now the Markov process $X_t = (x(t), J(t)) \in M = B \times \{1, 2\}$ induced by (20), where the rate matrix is given as

$$a = \begin{pmatrix} 0 & \tau(1-p) \\ \tau p & 0 \end{pmatrix} \tag{43}$$

with $0 < p < 1$ and $\tau > 0$. In other words, there is a Poisson clock with parameter τ and each time the clock rings, the process switches from its current environment to the other with probability p (respectively $1-p$) if the current environment is 2 (respectively 1).

Set

$$M_0^{bd} = \{(x,j) \in M : \ x_1x_2x_3 = 0\}, M_0^D = \{(x,j) \in M \ : x \in D\},$$

$$M_0 = M_0^{bd} \cup M_0^D,$$

and

$$M_+ = M \setminus M_0.$$

We shall prove here the following result.

Theorem 6.4 *Assume that*

$$p(\alpha_1 + \beta_1 - 2) + (1-p)(\alpha_2 + \beta_2 - 2) < 0$$

and

$$pr(x_*^1)\frac{\alpha_1 + \beta_1 - 2}{\alpha_1 + \beta_1 + 1} + (1-p)r(x_*^2)\frac{\alpha_2 + \beta_2 - 2}{\alpha_2 + \beta_2 + 1} > 0.$$

Then for τ sufficiently small, there is a unique persistent measure Π. Moreover

(i) Π *is absolutely continuous with respect to the Lebesgue measure* $dx_1dx_2dx_3 \otimes d(\delta_1 + \delta_2)$;

(ii) $Supp(\Pi) = \mathbf{C}(\Sigma^1, \Sigma^2) \times \{1, 2\}$;

(iii) *For all* $(x, i) \in M^+$,

$$\Pi_t \Rightarrow \Pi$$

$\mathbb{P}_{x,i}$ *almost surely;*

(iv) *Suppose r is constant on a open set meeting $\mathbf{C}(\Sigma^1, \Sigma^2)$; Then*

$$|P_t((x,i), \cdot) - \Pi| \leq Cst(1 + dist(x, \partial\mathbb{R}_+^3 \cup D)^{-\theta})e^{-\lambda t}$$

where θ, λ are positive constants (independent on (x,i)).

Remark 18 The assumption that r is constant on a open set meeting $\mathbf{C}(\Sigma^1, \Sigma^2)$ is an ad-hoc assumption chosen to simplify the computation of the Lie brackets involved in the verification of the strong bracket condition. We conjecture that the result holds true for any smooth r.

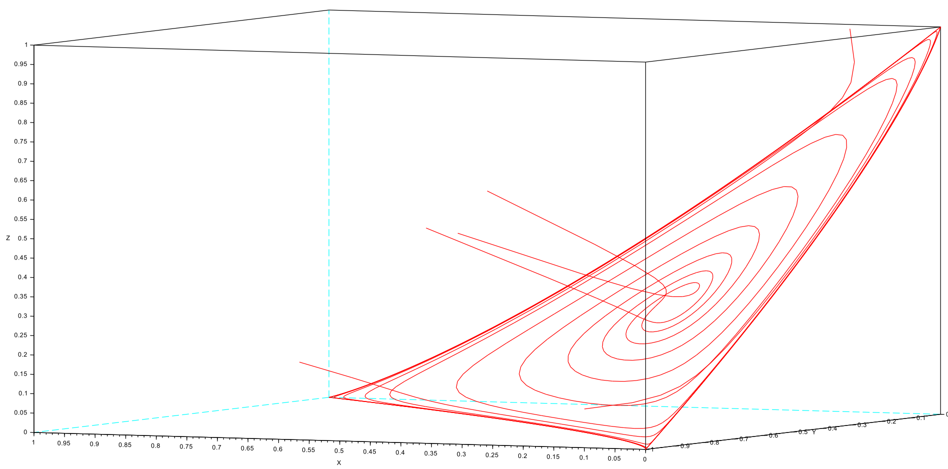

Figure 2: $\alpha_1 = 1,8; \beta_1 = 0,6$

Figures 2, 3, 4 result from simulations by Edouard Strickler and illustrate Theorem 6.4. Figures 2 and 3 picture the phase portraits of G^1, G^2 and Figure 4 is a realization of the switching process with $\tau = 10, p = 0,5$ and the function

$$r(x) = 100(\sum_{i=1}^{3} \exp[-200(x_i - z_*^1)^2])^{1/2}.$$

with $z_*^1 = (\alpha_1 + \beta_1 + 1)^{-1}$

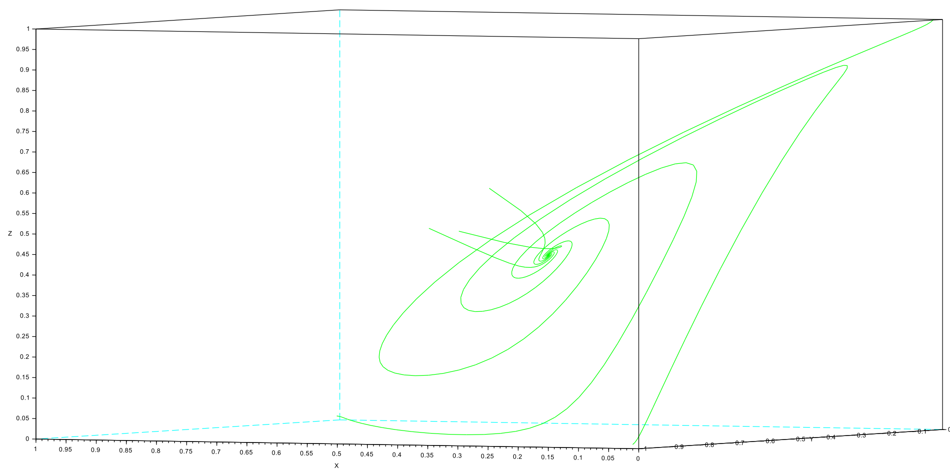

Figure 3: $\alpha_2 = 1,1; \beta_2 = 0,2$

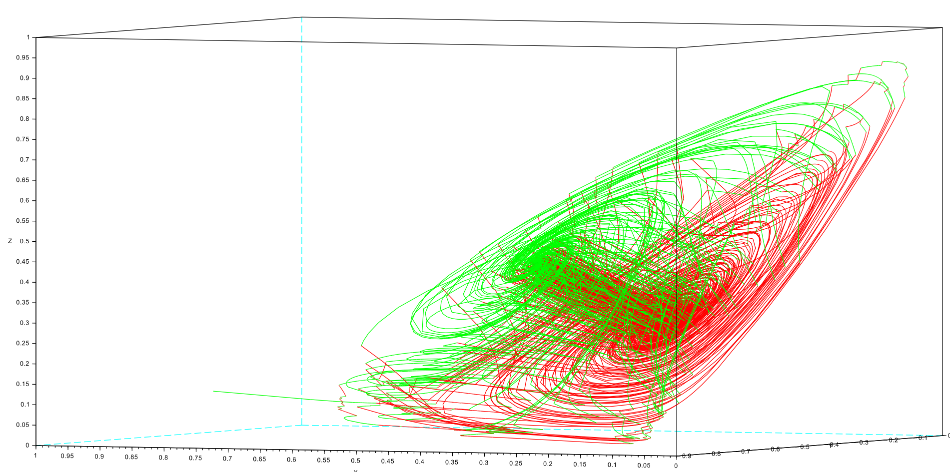

Figure 4: The switching process

Stochastic persistence with respect to M_0^{bd} The next proposition shows that permanence of the *average vector field* $pG^1 + (1-p)G^2$ implies stochastic persistence with respect to M_0^{bd}.

Proposition 6.5 *Let $\Lambda^{bd} = p(\alpha_1 + \beta_1 - 2) + (1-p)(\alpha_2 + \beta_2 - 2)$.*
If $\Lambda^{bd} < 0$, then (X_t) is H-persistent with respect to M_0^{bd}.

Proof: It is not hard to prove that on the face $x_i > 0, x_{i+1} = 0$, $x(t)$ converges to e_i. A (more general) proof can be found in [10], Theorem 3.1. Thus the only ergodic measures on M_0^{bd} are $\delta_{e_i} \otimes \nu, i = 1,2,3$ where ν is the Bernoulli measure on $\{1,2\}$ $\nu = p\delta_1 + (1-p)\delta_2$. The persistence criterion of Theorem 6.1 writes $\sum_j \nu_j(\alpha^j + \beta^j) < 2$ and the result follows. $\square$

Stochastic persistence with respect to M_0^D. Let $\ell : \mathbb{R}^3 \mapsto Ker(N) \times \mathbb{R}$ be the linear change of variable defined by $l(x) = (y,z)$ with

$$y = x - z(e_1 + e_2 + e_3) \text{ and } z = \frac{N(x)}{3}.$$

For $(y,z) \in Ker(N) \times \mathbb{R}$ set

$$(G_1^j(y,z), G_2^j(y,z)) = \ell \circ G^j \circ \ell^{-1}(y,z),$$

$\rho = \|y\|$ (the Euclidean norm of y), and $\theta = \frac{y}{\rho} \in S^1$ (the unit circle in $Ker(N)$) if $\rho \neq 0$.

In coordinates $(\rho, \theta, z) \in \mathbb{R}_+^* \times S^1 \times \mathbb{R}_+$, the dynamics of (X_t) in $B \setminus D \times \{1,2\}$ rewrites,

$$\left\{ \begin{array}{lcl} \dot{\rho} &=& \langle \theta, G_1^{J(t)}(\rho\theta, z) \rangle \\ \dot{\theta} &=& \frac{1}{\rho}[G_1^{J(t)}(\rho\theta, z) - \langle \theta, G_1^{J(t)}(\rho\theta, z) \rangle]. \\ \dot{z} &=& G_2^{J(t)}(\rho\theta, z) \end{array} \right. \tag{44}$$

This extends to a dynamics on

$$\tilde{M} = \{(\rho, \theta, z) : \ \ell^{-1}(\rho\theta, z) \in B\} \times \{1,2\},$$

leaving invariant the extinction set $\tilde{M}_0 = (\{0\} \times S^1 \times [\eta, 1]) \times \{1,2\}$, whose dynamics on $\tilde{M}_0$ is given by

$$\left\{ \begin{array}{lcl} \dot{\theta} &=& f^{J(t)}(\theta, z) \\ \dot{z} &=& g^{J(t)}(z) \end{array} \right. \tag{45}$$

where

$$f^j(\theta, z) = \partial_y G_1^j(0, z)\theta - \langle \theta, \partial_y G_1^j(0, z)\theta \rangle$$

and

$$g^j(z) = \hat{r}(z)z(1 - z(1 + \alpha_j + \beta_j)).$$

Here $\hat{r}(z)$ stands for $r(z(e_1 + e_2 + e_3))$.

Proposition 6.6 **(i)** *On $\tilde{M}_0$ the process $(\theta(t), z(t), J(t))$ (given by (45)) has a unique invariant measure $\mu = \mu^1(d\theta dz)\delta_1 + \mu^2(d\theta dz)\delta_2$.*

(ii) *Let*

$$\Lambda^D = \sum_{j=1,2} \int \langle \partial_y G_1^j(0, z)\theta, \theta \rangle \mu^i(d\theta dz).$$

If $\Lambda^D > 0$, then $(\rho(t), \theta(t), z(t), J(t))$ is H-persistent with respect to $\tilde{M}_0$.

(iii) *(slow and rapid switching). Assume the parameter p (in the definition of the rate matrix (43)) is fixed and write $\mu^j = \mu^j_\tau, \Lambda^D = \Lambda^D_\tau$ to emphasize the dependency on τ. Then*

(a)

$$\lim_{\tau \to 0} \Lambda^D_\tau = \frac{1}{2}[pr(x^1_*)\frac{\alpha_1 + \beta_1 - 2}{\alpha_1 + \beta_1 + 1} + (1 - p)r(x^2_*)\frac{\alpha_2 + \beta_2 - 2}{\alpha_2 + \beta_2 + 1}].$$

(b)

$$\lim_{\tau \to \infty} \Lambda^D_\tau = \frac{1}{2}[r(\bar{x}_*)\frac{\bar{\alpha} + \bar{\beta} - 2}{\bar{\alpha} + \bar{\beta} + 1}]$$

with

$$\bar{\alpha} = p\alpha_1 + (1 - p)\alpha_2, \bar{\beta} = p\beta_1 + (1 - p)\beta_2$$

and

$$\bar{x}_* = \frac{e_1 + e_2 + e_3}{1 + \bar{\alpha} + \bar{\beta}}.$$

Proof: *(i)* Let $U^j(\theta, z) = (f^j(\theta, z), g^j(z))$. We claim that for the dynamics induced by U^1, every point in $S^1 \times [\eta, 1]$ has $S^1 \times \{z^1_*\}$ as ω limit set and that $U^1(\theta, z^1_*)$ and $U^2(\theta, z^1_*)$ are linearly independent. This makes the point $(\theta, z^j_*, 1)$ an accessible point for the process $(\theta(t), \rho(t), J(t))$ at which the weak bracket condition is satisfied. The result follows from the standard arguments already used in the proof of Corollary 6.3.

We now prove the claim. It is easy to see that the Jacobian matrix $DG^j(x^j_*)$ leaves $Ker(N)$ invariant and that $DG^j(x^j_*)|_{Ker(N)}$, hence $\partial_y G^j_1(0, z^j_*)$, has two non real conjugates eigenvalues $\lambda_j, \overline{\lambda_j}$ with

$$\lambda_j = \frac{r(x^j_*)}{2}\left(\frac{\alpha_j + \beta_j - 2}{1 + \alpha_j + \beta_j} + i\sqrt{3}(\beta_j - \alpha_j)\right).$$

Therefore $\theta \mapsto f^j(\theta, z^j_*)$ never vanishes and the first part of the claim easily follows. For the second, note that $det(U^1(\theta, z^1_*), U^2(\theta, z^1_*)) = f^1(\theta, z^1_*)g^2(z^1_*) \neq 0$.

The H-persistence follows by choosing $V(\rho, \theta, z, j) = -\log(\rho)$ (for $0 < \rho \leq 1$) and

$$H((\rho, \theta, z, j)) = \begin{cases} \frac{1}{\rho}\langle \theta, G^j_1(\rho\theta, z)\rangle \text{ if } \rho \neq 0 \\ \langle \partial_y G^j_1(0, z)\theta, \theta\rangle \text{ if } \rho = 0 \end{cases}$$

(ii) The preceding discussion implies that U^j is uniquely ergodic on $S^1 \times [\eta, 1]$ with an invariant probability ν^j supported by $S^1 \times \{z^j_*\}$. Then

$$\int \langle \partial_y G^j_1(0, z)\theta, \theta\rangle \nu^j(d\theta dz) = \lim_{t\to\infty} \frac{\log(\| \exp{(t\partial_y G^j_1(0, z^j_*))}\|)}{t} = \Re(\lambda_j).$$

On the other hand, it is not hard to show that when p is fixed and $\tau \to 0$, every limit point of $\{\frac{\mu^1_\tau}{p}\}$ (respectively $\{\frac{\mu^2_\tau}{1-p}\}$) for the weak $*$ topology is invariant for U^1 (respectively U^2). Thus $\frac{\mu^1}{p} \Rightarrow \nu^1$, $\frac{\mu^2}{1-p} \Rightarrow \nu^2$, as $\tau \to 0$ and, consequently,

$$\lim_{\tau\to 0} \sum_{j=1,2} \int \langle \partial_y G^j_1(0, z)\theta, \theta\rangle \mu^j(d\theta dz) = p\Re(\lambda_1) + (1-p)\Re(\lambda_2).$$

For statement (b), remark that, by a standard averaging result, $\mu_\tau \Rightarrow \bar{\nu}$ as $\tau \to \infty$ where $\bar{\nu}$ is the invariant probability of the average vector field $\bar{U} = pU^1 + (1-p)U^2$. Thus, reasoning like in (a), $\sum_{j=1,2} \int \langle \partial_y G^j_1(0, z)\theta, \theta\rangle \mu^j_\tau(d\theta dz)$ converge, as $\tau \to \infty$, to the real part of the conjugate eigenvalues of $D\bar{G}(\bar{x}_*)|_{Ker(N)}$; Where $\bar{G} = pG^1 + (1-p)G^2$ and $\bar{x}_* = \frac{e_1+e_2+e_3}{1+\bar{\alpha}+\bar{\beta}}$ is the interior equilibrium of $\bar{G}$. $\square$

The accessible set We now characterize the accessible set $\mathbf{\Gamma}_{M^+}$.

Proposition 6.7 $\mathbf{\Gamma}_{M^+} = \mathbf{C}(\Sigma^1, \Sigma^2) \times \{1, 2\}$.

Proof: Relying on Proposition 6.2, we say that a point $p \in \mathbb{R}^3_+$ is (G^i)-accessible from $x \in \mathbb{R}^3_+$ if for every neighborhood O of p there exists a control u such that the solution $y(u, x, \cdot)$ to the control system (39) meets O. By Proposition 6.2, what we need to prove is that the set of points that are (G^i) accessible from any $x \in \mathbb{R}^3_+ \setminus (\partial\mathbb{R}^3_+ \cup D)$ coincide with $\mathbf{C}(\Sigma^1, \Sigma^2)$.

We first show that every $p \in \Delta$, is (G^i)-accessible from every $x \in \mathbb{R}^3_+ \setminus (D \cup \partial\mathbb{R}^3_+)$. We can always assume that $p \in \Delta \setminus (\partial\Delta \cup \{\frac{e_1+e_2+e_3}{3}\})$ because this latter set is dense in Δ.

Let $0 < s < 1$ be such that $s(\alpha_1 + \beta_1) + (1 - s)(\alpha_2 + \beta_2) = 2$ and let $\bar{G} = sG^1 + (1 - s)G^2$. Note that $\bar{G}$ is the vector field defined by (41) in the environment $(\bar{\alpha}, \bar{\beta}) = s(\alpha_1, \beta_1) + (1 - s)(\alpha_2, \beta_2)$.

Let $W(x) = \frac{x_1 x_2 x_3}{N^3(x)}$. A direct computation (see [42], Section 5.5) shows that W strictly decreases (respectively increases) along trajectories of G^1 (respectively G^2) in $\mathbb{R}^3_+ \setminus (D \cup \partial\mathbb{R}^3_+)$ and is constant along trajectories of $\bar{G}$.

If $W(x) > W(p)$ (respectively $<$) use the flow Φ^1, that is the control $u^1(t) = 1, u^2(t) = 0$ (respectively Φ^2) to steer x to a point x' at which $W(x') = W(p)$. Then use the flow $\bar{G}$, that is the control $u^1(t) = s, u^2(t) = 1 - s$, until $y(u, x', \cdot)$ meets O. Recall that Δ is a global attractor for $\bar{G}$ in $\mathbb{R}^3_+ \setminus \{0\}$ and that orbits on $\Delta \setminus \{\partial\Delta\}$ are periodic orbits (given as the level set of $W|_\Delta$).

We next show that every point $p \in \mathbf{C}(\Sigma^1, \Sigma^2)$ is (G^i)-accessible. Again it suffices to show that this is the case for $p \in \mathbf{C}(\Sigma^1, \Sigma^2) \setminus (\Sigma^1 \cup \Sigma^2)$. Such a point p lies in an interval $]\varsigma^1(q), \varsigma^2(q)[= \{t\varsigma^1(q) + (1 - t)\varsigma^2(q) \ : 0 < t < 1\}$ where $q \in \Delta \setminus \{e_1, e_2, e_3\}$. If $N(p) = 1$ $p \in \Delta$ and there is nothing to prove. If $N(p) > 1$, the characterization $\Sigma^2 = \partial_{\mathbb{R}^3_+} R^2(0)$ implies that $\lim_{t\to\infty} \Phi^2_{-t}(p) = 0$. Therefore $\Phi^2_{-t}(p) \in \Delta$ for some $t > 0$. Point $\Phi^2_{-t}(p)$ is then G^i accessible from x and so is p since $p = \Phi^2_t(\Phi^2_{-t}(p))$. If $N(p) < 1$ the proof is similar, using the characterization $\Sigma^1 = \partial_{\mathbb{R}^3_+} R^2(\infty)$. $\square$

Proof of Theorem 6.4 The following result implies Theorem 6.4. We use the notation of Propositions 6.5, 6.6, 6.7.

Theorem 6.8 **(i)** *If $\Lambda^{bd} > 0$ and $\Lambda^D > 0$, there exists a unique persistent measure Π verifying the conclusions of $(i),(ii),(iii)$ of Theorem 6.4 and conclusion (iv) for generic r.*

(ii) *If $\Lambda^{bd} < 0$ and $\Lambda^D > 0$, $x(t) \to \partial\mathbb{R}^3_+$ almost surely, for all $x(0) \in B \setminus D$;*

(iii) *If $\Lambda^{bd} > 0$ and $\Lambda^D < 0$, $x(t) \to D$ almost surely, for all $x(0) \in B \setminus \partial\mathbb{R}^3_+$;*

(iv) *If $\Lambda^{bd} < 0$ and $\Lambda^D < 0$. $x(t) \to \partial\mathbb{R}^3_+ \cup D$ almost surely for all $x(0) \in B \setminus (\partial\mathbb{R}^3_+ U D)$ and both events $x(t) \to \partial\mathbb{R}^3_+$ and $x(t) \to D$ have positive probability.*

Proof: We only prove the first assertion. The other ones are a consequence of the extinction results to be described in part II. They can also be proved directly like Theorems 3.1, 3.3 and 3.4 in [10].

In view of Propositions 6.5, 6.6, 6.7 and Theorem 4.10, it suffices to show that there exists a point $x \in \mathbf{C}(\Sigma^1, \Sigma^2)$ at which the weak (respectively strong) Hörmander condition is satisfied. Let

$$C^i = \begin{bmatrix} -1 & -\alpha_i & -\beta_i \\ -\beta_i & -1 & -\alpha_i \\ -\alpha_i & -\beta_i & -1 \end{bmatrix}.$$

For $x \in \mathbb{R}^3_+$ let $\mathsf{diag}(x)$ denote the diagonal matrix whose entries are the components of x and let $\mathbf{1} = e_1 + e_2 + e_3$. Then $G^i(x) = r(x)U^i(x)$ with $U^i(x) = \mathsf{diag}(x)(\mathbf{1} + C^i x)$. Since the term $r(x)$ has no incidence on the weak bracket condition, it suffices to verify that it holds for the vector fields U^1, U^2. A straightforward computation show that

$$[U^2, U^1](x) = \mathsf{diag}(x)(C^1\mathsf{diag}(x)C^2x - C^2\mathsf{diag}(x)C^1x) + U^1(x) - U^2(x).$$

Thus $Det(U^1(x), U^2(x), [U^2, U^1](x)) = (x_1x_2x_3)^3 P(x)$ where

$$P(x) = Det(\mathbf{1} + C^1x, \mathbf{1} + C^2x, C^1\mathsf{diag}(x)C^2x - C^2\mathsf{diag}(x)C^1x).$$

Since the function P is a polynomial in the variables x_1, x_2, x_3, it suffices to show that is is not identically 0 to deduce that $P(x) \neq 0$ for some x in the interior of $\mathbf{C}(\Sigma^1, \Sigma^2)$. The tedious computation of the coefficients of P becomes a child's play with the help of the mathematical software Python/Sympy and the great help of Jean Baptiste Bardet who knows how to use it. It appears that the coefficient of the monomial $x_1x_2x_3$ is

$$P_{1,1,1} = 3\left[(\alpha_1 + \beta_1 - 2)(\beta_2 - 1) - (\beta_1 - 1)(\alpha_2 - \beta_2 - 2)\right](\alpha_1 + \beta_1 - (\alpha_2 + \beta_2))$$

which is never 0. This concludes the proof of the weak bracket condition.

For the strong bracket condition, under our assumption that r is constant on a open set meeting $\mathbf{C}(\Sigma^1, \Sigma^2)$, it suffices to show that

$$Q(x) = \frac{1}{(x_1x_2x_3)^3} Det(U^1(x) - U^2(x), [U^2, U^1](x), [[U^2, U^1], U^1](x))$$

is a non zero polynomial. Thanks again to Python/Sympy and Jean Baptiste Bardet, the coefficient of $x_1^4x_2^2$ in Q is

$$Q_{4,2,0} = \alpha_1^2\beta_2^2 - 2\alpha_1^2\beta_2 - 2\alpha_1\alpha_2\beta_1\beta_2 + 2\alpha_1\alpha_2\beta_1 + 2\alpha_1\alpha_2\beta_2 - \alpha_1\beta_1\beta_2$$

$$+2\alpha_1\beta_1 + \alpha_1\beta_2^2 - 2\alpha_1\beta_2 + \alpha_2^2\beta_1^2 - 2\alpha_2^2\beta_1 + \alpha_2\beta_1^2 - \alpha_2\beta_1\beta_2 - 2\alpha_2\beta_1 + 2\alpha_2\beta_2.$$

The coefficient of $x_1^2x_2^3$ is

$$Q_{2,3,0} = -2\alpha_1^3\beta_2 + 2\alpha_1^2\alpha_2\beta_1 + 2\alpha_1^2\alpha_2\beta_2 - 2\alpha_1^2\beta_1 + 2\alpha_1^2\beta_2 - 2\alpha_1\alpha_2^2\beta_1$$

$$+2\alpha_1\alpha_2\beta_1 - 2\alpha_1\alpha_2\beta_2 + 2\alpha_1\beta_1^2\beta_2 - 2\alpha_1\beta_1^2 - 2\alpha_1\beta_1\beta_2^2 + 2\alpha_1\beta_1\beta_2 - 4\alpha_1\beta_1 + 4\alpha_1\beta_2 - 2\alpha_2\beta_1^3 + 2\alpha_2\beta_1^2\beta_2$$

$$+2\alpha_2\beta_1^2 - 2\alpha_2\beta_1\beta_2 + 4\alpha_2\beta_1 - 4\alpha_2\beta_2$$

The solutions of the polynomial equation $Q_{4,2,0} = Q_{2,3,0} = 0$ are the sets

$$\{\alpha_1 = -1, \beta_1 = 0\}, \{\alpha_1 = 0, \alpha_2 = 0\}, \{\alpha_1 = 0, \beta_1 = 2\}, \{\alpha_1 = -\beta_1 - 1, \alpha_2 = -\beta_2 - 1\},$$

$$\{\alpha_1 = -\beta_1 + 2, \alpha_2 = -\beta_2 + 2\}, \{\alpha_1 = \beta_1/2 - 1, \alpha_2 = \beta_2/2 - 1\}, \{\beta_1 = 0, \beta_2 = 0\}.$$

None of these solutions is compatible with the constraints on the parameters. Hence Q is non zero and the strong bracket condition holds true. $\square$

7 Proof of Theorem 4.4

Since M is locally compact and separable there exists a sequence $\{C_n\}_{n\geq 1}$ of compact sets with $C_n \subset \operatorname{int}(C_{n+1})$ such that $M = \cup_{n\geq 1} C_n$. Throughout we let

$$K_n = \{x \in M : d(x, M_0) \geq \frac{1}{n}\} \cap C_n. \tag{46}$$

Note that

$$M_+ = \bigcup_{n\geq 1} K_n$$

and

$$M_0 = \bigcap_{n\geq 1} K_n^c = \bigcap_{n\geq 1} \overline{K_n^c}$$

where the later equality follows from the inclusion $C_n \subset \mathrm{int}(C_{n+1})$.

For any function $f \in \mathcal{D}(\mathcal{L})$ it is well known (see e.g [25] or [50]) that the process

$$M_t^f(x) \stackrel{def}{=} f(X_t^x) - f(x) - \int_0^t (\mathcal{L}f)(X_s^x)ds,\ t \geq 0. \tag{47}$$

is a $(\mathcal{F}_t)$ Martingale. We let $\langle M^f(x)\rangle_t$ denote its *predictable quadratic variation*, defined as the compensator of $(M_t^f(x))^2$.

Lemma 7.1 *Let $f \in \mathcal{D}^2(\mathcal{L})$. Then*

$$\langle M^f(x)\rangle_t = \int_0^t (\Gamma f)(X_s^x)ds \tag{48}$$

Proof: The map $t \to \int_0^t (\Gamma f)(X_s^x)ds$ is nondecreasing and continuous (hence predictable). It then remains to show that $((M_t^f)^2 - \int_0^t (\Gamma f)(X_s^x)ds)$ is a Martingale. This is a folklore result, for which we provide a proof. Let $M_t = f(x) + M_t^f(x)$ and $N_t = f^2(x) + M_t^{f^2}(x)$. Then $\{M_t\}_{t\geq 0}$ and $(N_t)_{t\geq 0}$ are both martingales. It then suffices to prove that (Z_t) is a martingale, where $Z_t = M_t^2 - \int_0^t \Gamma f(X_s^x)ds - N_t$. Set $g_t = \mathcal{L}f(X_t^x)$ and $G_t = \int_0^t g_s ds$. Then

$$Z_t = (f(X_t^x) - G_t)^2 - \int_0^t (\Gamma f)(X_s^x)ds - (f^2(X_t^x) - \int_0^t \mathcal{L}(f^2)(X_s^x)ds)$$

$$= 2\int_0^t f(X_s^x)g_s ds + G_t^2 - 2f(X_t)G_t = 2\int_0^t (G_s + M_s)g_s ds + G_t^2 - 2(G_t + M_t)G_t.$$

By Fubini formulae $G_t^2 = 2\int_0^t G_s g_s ds$. Thus $Z_t = 2\int_0^t M_s g_s - G_t M_t$ and

$$Z_{t+u} - Z_t = 2\int_t^{t+u} (M_s - M_{t+u})g_s ds + (M_t - M_{t+u})G_t.$$

From this expression it is clear that $\mathsf{E}(Z_{t+u} - Z_t|\mathcal{F}_t) = 0$ for all $t, u \geq 0$. $\square$

Lemma 7.2 *Let τ be a stopping time. Then*

$$\mathbf{1}_{M_0}(X_\tau^x) = \mathbf{1}_{M_0}(x)\ a.s\ on\ \tau < \infty.$$

Proof: If $x \in M_0$ the event $E = \cap_{t \in \mathbb{Q}^+} \{X_t^x \in M_0\}$ has probability one by Hypothesis 1. By right continuity of paths and closeness of M_0 $E \subset \cap_{t \in \mathbb{R}^+} \{X_t^x \in M_0\}$. In particular $X_\tau^x \in M_0$ a.s on $\{\tau < \infty\}$.

Suppose now $x \in M_+$. Let $f \in \mathcal{D}(\mathcal{L})$. For all $n, N \in \mathbb{N}^*$

$$E(f(X_N^x)\mathbf{1}_{\{\tau \leq N\}}) = \sum_{i=1}^{2^n N} \mathsf{E}(f(X_N^x)\mathbf{1}_{\{(i-1)2^{-n} < \tau \leq i2^{-n}\}})$$

$$= \sum_{i=1}^{2^n N} \mathsf{E}(\mathsf{E}(f(X_N^x)|\mathcal{F}_{i2^{-n}})\mathbf{1}_{\{(i-1)2^{-n} < \tau \leq i2^{-n}\}})$$

$$= \sum_{i=1}^{2^n N} \mathsf{E}(P_{N-i2^{-n}} f(X_{i2^{-n}}^x)\mathbf{1}_{\{(i-1)2^{-n} < \tau \leq i2^{-n}\}}) = \mathsf{E}(P_{N-\tau(n)} f(X_{\tau(n)}^x)\mathbf{1}_{\tau \leq N})$$

where

$$\tau(n) = \sum_{i=1}^{2^n N} i2^{-n}\mathbf{1}_{\{(i-1)2^{-n} < \tau \leq i2^{-n}\}}.$$

Since $f \in \mathcal{D}(\mathcal{L})$ $\|P_t f - f\| \to 0$ as $t \to 0$. Thus, by $C_b(M)$ Feller continuity (Hypothesis 2) and right continuity of $t \to X_t^x$,

$$\lim_{n \to \infty} P_{N-\tau(n)} f(X_{\tau(n)}^x) = P_{N-\tau} f(X_\tau^x),$$

a.s on $\tau \leq N$. It then follows (by dominated convergence) that

$$E(f(X_N^x)\mathbf{1}_{\{\tau \leq N\}}) = \mathsf{E}(P_{N-\tau} f(X_\tau^x)\mathbf{1}_{\{\tau \leq N\}}). \tag{49}$$

Let now $f \in C_b(M)$. Then $f_n := \int_0^1 P_{s/n} f ds \in \mathcal{D}(\mathcal{L}), \lim_{n \to \infty} f_n = f$ pointwise and $\|f_n\| \leq \|f\|$ (details are left to the reader). Therefore, property (49) holds also for f (use dominated convergence again, once on the left hand side and twice on the right hand side of (49)).

Let now $f_n \in C_b(M)$ be such that $f_n \downarrow \mathbf{1}_{M_0}$. For instance $f_n(x) = (1 - nd(x, M_0))_+$. Then, by monotone convergence, (49) holds with $f = \mathbf{1}_{M_0}$. Thus

$$\mathsf{E}(\mathbf{1}_{M_0}(X_\tau^x)\mathbf{1}_{\{\tau \leq N\}}) = \mathsf{E}(P_{N-\tau}\mathbf{1}_{M_0}(X_\tau^x)\mathbf{1}_{\{\tau \leq N\}})$$

$$= \mathsf{E}(\mathbf{1}_{M_0}(X_N^x)\mathbf{1}_{\{\tau \leq N\}}) \leq P_N \mathbf{1}_{M_0}(x) = 0$$

where the first and last inequalities follows from Hypothesis 1 and the second from (49). $\square$

Lemma 7.3 *Let $x \in M_+$ and $\tau_n(x) \stackrel{def}{=} \inf\{t \geq 0 : X_t^x \in K_n^c\}$. Then $\{\tau_n(x)\}_{n\geq 1}$ is a localizing sequence. That is $\tau_n(x)$ is a stopping time and* $\lim_{n\to\infty} \tau_n(x) = \infty$.

Proof: Fix $x \in M_+$ and set $\tau_n = \tau_n(x)$. Then τ_n is stopping time as K_n^c is open and the filtration right continuous. Obviously, $\tau_n \leq \tau_{n+1}$. Hence, $\lim_{n\to\infty} \tau_n = \tau \in \mathbb{R}^+ \cup \{\infty\}$ exists a.s. and is a stopping time. Furthermore, by Lemma 7.2, $\tau_n < \tau$ a.s. on $\tau < \infty$ (since on $\{\tau_n = \tau;\, \tau < \infty\}$ $X_{\tau_n}^x = X_\tau^x \in \cap_{m\geq n} K_m^c = M_0$). The fact that $\tau_n < \tau$ implies that (X_t) is almost surely left continuous at τ (i.e $X_{\tau^-} = X_\tau$) on $\tau < \infty$. This later property knows as a the quasi left continuity property is often proved for Feller processes but the proof only requires the cad-lag continuity of paths and the strong Markov property (see Remark 2). Since $X_{\tau^-}^x \in M_0$ on $\tau < \infty$ we get that $X_\tau^x \in M_0$ and the conclusion follows from Lemma 7.2. □

Lemma 7.4 *Assume Hypothesis 4. Then for all $x \in M_+$ the process $\{M_t^V(x)\}_{t\geq 0}$ defined by*

$$M_t^V(x) = V(X_t^x) - V(x) - \int_0^t H(X_s^x)ds,\ t \geq 0, \tag{50}$$

is a square integrable martingale and

$$\lim_{t\to\infty} \frac{M_t^V(x)}{t} = 0, \text{-}a.s. \tag{51}$$

Proof: For all $x \in M_+, t \to X_t^x \in M_+$ and has cadlag paths. Thus $\{M_t^V(x)\}_{t\geq 0}$ is well-defined . Let $\{K_n\}_{n\geq 1}$ be as defined in (46) and $\{\tau_n(x)\}_{n\geq 1}$ be as defined in Lemma 7.3. Set $\tau_n = \tau_n(x)$. Then, by Hypothesis 4 $(i),(a)$ and Lemma 7.3,

$$M_{t\wedge\tau_n}^V(x) = V(X_{t\wedge\tau_n}^x) - V(x) - \int_0^{t\wedge\tau_n} H(X_s^x)ds = V_{K_n}(X_{t\wedge\tau_n}^x) - V(x) - \int_0^{t\wedge\tau_n} (\mathcal{L}V_{K_n})(X_s^x)ds$$

is a martingale. Then, $\{M_t^V(x)\}_{t\geq 0}$ is a local martingale. Now, by Lemma 7.1 and Hypothesis 4 $(i),(b)$

$$\mathsf{E}(\langle M^V(x)\rangle_{t\wedge\tau_n}) = \mathsf{E}(\int_0^{t\wedge\tau_n} (\Gamma V_{K_n})(X_s^x)ds) \leq \int_0^t P_s\Gamma(V_{K_n})(x)ds \leq C_x t$$

for some constant C_x. Hence,

$$\mathsf{E}(\langle M^V(x)\rangle_t) \leq C_x t < \infty \tag{52}$$

This makes $\{M_t^V(x)\}$ a (true) L^2 martingale and $\{(M_t^V)^2 - \langle M^V(x)\rangle_t\}$ a martingale. A proof can be found in [50], theorem 4.3 for continuous martingales. The proof extends verbatim for right continuous martingales (provided we replace the quadratic variation by the predictable quadratic variation).

For the last part of the Lemma, let $m_t = \int_0^t \frac{1}{1+s} dM_s^V(x)$. Then $\{m_t\}$ is a local martingale and

$$\mathsf{E}(\langle m\rangle_\infty) = \mathsf{E}(\int_0^\infty \frac{1}{(1+s)^2} d\langle M_s^V\rangle) \leq \int_0^\infty C_x \frac{1}{(1+s)^2} ds.$$

Hence, by the strong law of large number for local martingales (see Lipster [53], Theorem 1), $\lim_{t\to\infty} \frac{M_t^V(x)}{t} = 0.a.s.$ □

The proof of the following Lemma is similar to the proof of Proposition 1 in [70].

Lemma 7.5 *Assume that* $\{(X_t^x)_{t\geq 0} : x \in M_+\}$ *is* H*-persistent. Then*

(i) For all $\mu \in \mathcal{P}_{inv}(M)$, $\mu H \leq 0$; *and* $\mu H = 0 \Leftrightarrow \mu \in \mathcal{P}_{inv}(M_+)$.

(ii) $\mathcal{P}_{inv}(M_+)$ *is tight :* $\forall \varepsilon > 0, \exists K \subset M_+$ *compact such that* $\inf\{\mu(K) : \mu \in \mathcal{P}_{inv}(M_+)\} \geq 1 - \varepsilon$.

Proof: (i). By Hypothesis 4 (ii), Proposition 2.1 and Lemma 9.1 (ii), $H \in L^1(\mu)$ for all $\mu \in \mathcal{P}_{inv}(M)$. Let $\mu \in \mathcal{P}_{inv}(M_+)$. We claim that $\mu H = 0$. By the ergodic decomposition theorem it suffices to prove the result for μ ergodic. By Birkhoff ergodic Theorem, for μ almost all x and $\mathbb{P}_x$ almost surely

$$\lim_{t\to\infty} \Pi_t^x H = \mu H.$$

Hence, by Proposition 7.4,

$$\lim_{t\to\infty} \frac{V(X_t^x)}{t} = \mu H.$$

Since $\mu(M_0) = 0$ there exists $n \geq 1$ such that $\mu(K_n) \geq 1/2$, so that, by Birkhoff ergodic Theorem again, $(X_t^x)_{t\geq 0}$ visits K_n infinitely often for μ almost all x, $\mathbb{P}_x$ almost surely. Since V is bounded on K_n this proves that $\mu H = 0$.

Let now $\mu \in \mathcal{P}_{inv}(M) \setminus \mathcal{P}_{inv}(M_+)$. We can write, by Hypothesis 1, $\mu = (1-t)\mu_0 + t\mu_1,\ 0 \le t < 1$, with $\mu_0 \in \mathcal{P}_{inv}(M_0)$ and $\mu_1 \in \mathcal{P}_{inv}(M_+)$. Thus $\mu H = (1-t)\mu_0 H < 0$

(ii). Suppose not. Then there exists some $\epsilon > 0$ such that for each $n \ge 1$ there exists some $\mu_n \in \mathcal{P}_{inv}(M_+)$ with $\mu_n(K_n) < 1-\epsilon$. Thus, $\mu_n(\overline{K_m^c}) > \epsilon$ for all $m < n$ as, by definition, $\overline{K_{n+1}^c} \subset \overline{K_n^c}$. Let μ be a limit point of (μ_n) for the weak* topology. Then $\mu \in \mathcal{P}_{inv}(M)$ as $\mathcal{P}_{inv}(M)$ is tight and, by application of Portemanteau $\mu(\overline{K_m^c}) \ge \epsilon$ for all $m \ge 1$. Since $M_0 = \cap_{m\ge1}\overline{K_m^c}$ this implies $\mu(M_0) \ge \epsilon$. Now, by by part (i) $\mu_n H = 0$ implying $\mu H = 0$ and $\mu \in \mathcal{P}_{inv}(M_+)$ again by part (i). A contradiction. $\square$

Remark 19 The proof of Lemma 7.5 (i) also shows that when $M_0 = \emptyset$, then $\mu H = 0$ for all $\mu \in \mathcal{P}_{inv}(M)$.

We now prove Theorem 4.4.

(i). By Proposition 2.1, for every $x \in M_+$, every weak limit point of $(\Pi_t^x)_{t\ge0}$ lies $\mathbb{P}_x$-a.s. in $\mathcal{P}_{inv}(M)$. Let $\mu = \lim_{n\to\infty} \Pi_{t_n}^x \in \mathcal{P}_{inv}(M)$ be such a weak limit point. Then, by Proposition 7.4 again, $\lim_{n\to\infty} \frac{V(X_{t_n}^x)}{t_n} = \mu H \ge 0$ (because $V \ge 0$) and the result follows from assertion (i) of Lemma 7.5.

(ii). Suppose that this is not true implying that there exists some $\varepsilon > 0$ and a sequence $(x_n) \subset M_+$ such that

$$\mathsf{P}(\liminf_{t\to\infty} \Pi_t^{x_n}(K_n^c) \ge \varepsilon) > 0.$$

By assertion (i), with probability 1 there exists a subsequence $\{t_m^n\}_{m\ge1} \uparrow \infty$ and $\mu_n \in \mathcal{P}_{inv}(M_+)$ such that $\liminf_{t\to\infty} \Pi_t^{x_n}(K_n^c) = \lim_{m\to\infty} \Pi_{t_m^n}^{x_n}(K_n^c)$ and $\Pi_{t_m^n}^{x_n} \Rightarrow \mu_n$ as $m \to \infty$. Thus, by Portemanteau theorem, $\mu_n(\overline{K_n^c}) \ge \varepsilon$ on the event $\liminf_{t\to\infty} \Pi_t^{x_n}(K_n^c) \ge \varepsilon$. By tightness of $\mathcal{P}_{inv}(M_+)$, (Lemma 7.5, (ii)) for n large enough $\mu_n(\overline{K_n^c}) < \varepsilon$. A contradiction.

8 Return times and convergence rates

Theorem 4.4 shows that, under H persistence, the process (X_t^x) spends most of its time in a compact set far from the extinction set. Here we are interested in more quantitative consequences of H persistence. First we will estimate the mean time needed to reach such a compact set.Then we will give condition ensuring that the rate of convergence in Theorem """ is exponential Throughout the remainder of this section we assume, that the process $\{X_t^x :\ x \in M_+\}$ is H-persistent, that is $\Lambda^-(H) > 0$,

The case M_0 compact

We assume here that M_0 is compact. For $\delta > 0$ we let

$$M_0^\delta = \{x \in M_+ :\ d(x, M_0) < \delta\}.$$

Proposition 8.1 *Let $0 < \lambda < \Lambda^-(H)$. For every $T_0 > 0$ (sufficiently large) and $T_1 > T_0$, there exists $\delta > 0$ such that for all $T \in [T_0, T_1]$*

$$P_T V(x) - V(x) \quad \leq \quad -\lambda T \text{ for } x \in M_0^\delta.$$

Given such a $T \in [T_0, T_1]$, let

$$\tau_1 \stackrel{def}{=} \inf\{k \in \mathbb{N}^* : X_{kT}^x \in M_+ \setminus M_0^\delta\}$$

and

$$\tau_n = \inf\{k \in \mathbb{N} : k > \tau_{n-1}, X_{nT}^x \in M_+ \setminus M_0^\delta\}.$$

Then

$$\mathsf{E}_x(\tau_1) \leq \begin{cases} \dfrac{V(x)}{\lambda T} \text{ if } x \in M_0^\delta, \\ \\ 1 + \dfrac{P_T V(x)}{\lambda T} \text{ if } x \in M_+ \setminus M_0^\delta \end{cases}$$

and

$$\mathsf{E}_x(\tau_{n+1}) \leq \mathsf{E}_x(\tau_n) + 1 + \frac{v(\delta, T)}{\lambda T}$$

for all $n \geq 1$, where

$$v(\delta, T) := \sup\{P_T V(x) :\ x \in M_+ \setminus M_0^\delta\}.$$

Proof: For all $x \in M, t \geq 0$ we let

$$\overline{H}(t, x) = \int_0^t P_s H(x).$$

Recall, that by Lemma 7.4,

$$P_T V(x) - V(x) = \overline{H}(T, x)$$

for all $x \in M_+, T \geq 0$. The first assertion then follows from the two following facts (and compactness of M_0):

(a) There exists $T_0 > 0$ (arbitrary large) such that all $x \in M_0$ and $t \geq T_0$ $\overline{H}(t, x) < -\lambda t$;

(b) $\overline{H}$ is continuous in (t, x).

Proof of a). Suppose the contrary. Then for all $n \in \mathbb{N}^*, \exists t_n \geq n, x_n \in M_0$ such that $\mu_n H \geq -\lambda$ where μ_n stands for the measure defined by

$$\mu_n f = \frac{1}{t_n} \int_0^{t_n} P_s f(x_n) ds$$

for all $f \in \mathcal{M}_b(M)$. By Proposition 2.1, $\mu_n \tilde{W} \leq \frac{W(x_n)}{t_n} + C$. Thus, by Lemma 9.1, (μ_n) is tight. Let μ be a limit point of (μ_n) it is easily seen that $\mu \in \mathcal{P}_{inv}(M_0)$ (because for all $f \in C_b(M_0), |\mu_n f - \mu_n P_t f| \leq \frac{2t\|f\|}{t_n} \to 0$ as $n \to \infty$). By Hypotheses 4 (ii), 3 and Lemma 9.1 (ii) we get that $\mu H \geq -\lambda > -\Lambda^-(H)$. A contradiction.

Proof of b). Let (μ_n) be defined like in the proof of (a) but, this time with $t_n \to t^*$ and $x_n \to x_*$. The sequence (μ_n) is tight, and for every limit point ν of (μ_n) and all $f \in C_b(M)$ $\nu f = \frac{1}{t^*} \int_0^{t*} P_s f(x_*) ds$ by $C_b(M)$ Feller continuity. Thus $\nu f = \frac{1}{t^*} \int_0^{t*} P_s f(x^*) ds$ for all $f \in L^1(\nu)$. In particular $\lim_{n \to \infty} \overline{H}(t_n, x_n) = \overline{H}(t, x)$.

The last assertions follow from Pakes's criterion (see Theorem 9.1.2 and its proof in [22]) □

Remark 20 Although there is no evidence that the quantity $v(\delta, T)$ in Proposition 8.1 is finite, we can always modify V and H outside a neighborhood of M_0 such that:

(i) H is bounded on M, and

(ii) V is bounded on $M_+ \setminus M_0^\delta$ for all $\delta > 0$;

In particular

$$v(\delta, T) \leq \sup\{V(x) \ : x \in M_+ \setminus M_0^\delta\} + T\|H\|.$$

Indeed, let C be a compact set such that $M_0 \cup M_0^{\delta_0} \subset \mathrm{int}(C)$ for some $\delta_0 > 0$. The set $K = C \setminus (M_0 \cup M_0^{\delta_0})$ is a compact subset of M_+. Set $\tilde{V}(x) = V(x)$ if $x \in C \setminus M_0$, $\tilde{H}(x) = H(x)$ if $x \in C$, $\tilde{V}(x) = V_K(x)$ and $\tilde{H}(x) = \mathcal{L}(V_K)(x)$ for $x \notin C$. The map $(\tilde{V}, \tilde{H})$ coincide with (V, H) on $C \setminus M_0 \times C$ and satisfies the required conditions.

Proposition 8.2 *Assume that the process is H-persistent (strong version) with V and H like in Remark 20. Then, for every $T_0 > 0$ (sufficently large) and $T_1 > T_0$, there exist positive numbers θ, δ, κ and $\rho < 1$ such that for all $T \in [T_0, T_1]$*

$$P_T(e^{\theta V})(x) \leq \begin{cases} \rho e^{\theta V(x)} \ \textit{on } M_0^\delta, \\ \\ \kappa \ \textit{on } M_+ \setminus M_0^\delta \end{cases}$$

Furthermore, letting $b = 1/\rho$,

$$\mathsf{E}_x(b^\tau) \leq \begin{cases} e^{\theta V(x)} \ \textit{if } x \in M_0^\delta, \\ \\ b(1 + P_T e^{\theta V}(x)) \ \textit{if } x \in M_+ \setminus M_0^\delta \end{cases}$$

and

$$\mathsf{E}_x(b^{\tau_{n+1}}) \leq \mathsf{E}_x(b^{\tau_n}) b(1 + \kappa) \tag{53}$$

for $n \geq 1$.

Proof: Let $\lambda, T_0, T_1, \delta$ be as in Proposition 8.1. For $x \in M_+$ and $T \in [T_0, T_1]$

$$e^{\theta V(X_T)} = e^{\theta V(x)} e^{\theta(P_T V(x) - V(x))} e^{\theta(M_T^1(x) + M_T^V(x))} \tag{54}$$

where

$$M_T^1(x) = \int_0^T H(X_s^x) - P_s H(x) ds$$

and $(M_t^V(x))$ is the Martingale defined in Lemma 7.4.

Observe that $\mathsf{E}(M_T^1(x)) = 0$ and $M_T^1(x)^2 \leq c_1 T^2$ with $c_1 = 4\|H\|^2$. Thus, by elementary properties of the log-laplace transform

$$\mathsf{E}(e^{\theta M_T^1(x)}) \leq e^{c_1 T^2 \frac{\theta^2}{2}}. \tag{55}$$

Let

$$r(x) = e^x - x - 1 \leq x^2 e^x.$$

Using Lemma 26.19 of [48], the process

$$Z_t(\theta) = \exp\big(\theta M_t^V(x) - \frac{r(\theta \Delta V)}{(\Delta V)^2} \langle M_t^V(x) \rangle\big) \tag{56}$$

is a supermartingale for all $\theta > 0$ (here and below we adopt the convention that $\frac{r(\theta\Delta V)}{(\Delta V)^2} = \frac{\theta^2}{2}$ when $\Delta V = 0$). Thus,

$$\mathsf{E}(\exp\big(\theta M_T^V(x) - \frac{r(\theta \Delta V)}{(\Delta V)^2} \gamma T)\big) \leq 1 \tag{57}$$

where γ is the supremum in condition (ii). It follows, by Hölder inequality, that

$$\mathsf{E}(e^{\theta(M_T^1(x) + M_T^V(x))}) \leq e^{c_1 T^2 \theta^2} e^{\gamma T \frac{r(2\theta\Delta V)}{2(\Delta V)^2}} \leq e^{c_2 \theta^2 T}$$

with $c_2 = c_1 T_1 + 2\gamma e$ and $2\theta\Delta V \leq 1$. This latter inequality combined with (54) and Proposition 8.1 proves the first assertion with

$$\rho = e^{-\theta T_0(\lambda - c_2\theta)}, \kappa = e^{\theta\left(T_1 \|H\| + \sup\{V(x) \colon x \in M_+ \setminus M_0^\delta\} + c_2 \theta T_1\right)}$$

and θ small enough. The last assertion follows by observing that

$$W_n = e^{\theta V(X_{n\wedge\tau})} b^{n \wedge \tau}$$

is a supermartingale with respect to $(\mathcal{F}_{nT})$. Hence, for $x \in M_0^\delta$

$$\mathsf{E}_x(b^{n\wedge\tau}) \leq \mathsf{E}_x(W_n) \leq W_0 = e^{\theta V(x)}$$

while for $x \notin M_0^\delta$

$$\mathsf{E}_x(b^\tau) = b\mathsf{E}_x(\mathbf{1}_{X_T \notin M_0^\delta} + \mathsf{E}_{X_T}(b^\tau)\mathbf{1}_{X_T \in M_0^\delta}) \leq b(1 + P_T e^{\theta V}(x))$$

by the Markov property (compare to the proof of Theorem 8.1.5(2) of [22]) The bound for $\mathsf{E}_x(b^{\tau_n})$ is obtained similarly by using the strong Markov property. $\square$

8.1 The case M_0 non compact

The purpose of this section is to prove the following result, similar to Proposition 8.2, when M_0 is noncompact. The proof is inspired by the proof of Proposition 4.1 in [37].

Proposition 8.3 *Assume that the process is H-persistent (strong version') (see Hypothesis 5 and condition (ii)' in Hypothesis 4) and* persistent at infinity*, meaning that V is proper and there exists a compact $C \subset M$ such that*

$$\sup_{x \in M \setminus C} H(x) < 0.$$

Then there exist positive numbers $T_1 > T_0$, θ, κ and $\rho < 1$ such that for all $T \in [T_0, T_1]$

$$P_T(e^{\theta V})(x) \leq \rho e^{\theta V(x)} + \kappa.$$

From now on and throughout the section we assume that the process is $H-$ persistent (strong version) and without loss of generality[5], that

$$H(x) \leq -2$$

on $M \setminus C$.

We let θ_0 denote a positive number small enough so that

$$\gamma \theta_0 e^{\theta_0 \Delta V} \leq 1$$

where $\gamma, \Delta V$ are like in Hypothesis 4'.

Lemma 8.4 *There exists $\omega_0 > 0$ such that for all $\theta \leq \theta_0, T \geq 0$ and $x \in M_+$*

$$P_T e^{\theta V}(x) \leq e^{\theta V(x)} e^{\theta \omega_0 T},$$

and

$$\mathsf{E}_x(e^{\theta(V(T \wedge \tau) + T \wedge \tau)}) \leq e^{\theta V(x)}.$$

where $\tau = \inf\{t \geq 0 : \; X_t^x \in C\}$;

Proof: Let

$$H^* = \sup\{|H(x)| : \; x \in C\}. \tag{58}$$

[5] It suffices to multiply V and H by a sufficient large constant

By persistence at infinity, $H(x) \leq H^*$ for all $x \in M$. From (50)

$$P_T e^{\theta V}(x) \leq e^{\theta V(x)} e^{\theta H^* T} \mathsf{E}(e^{\theta M_T^V(x)}).$$

By (57)

$$\mathsf{E}(e^{\theta M_T^V(x)}) \leq e^{\gamma T \frac{r(\theta \Delta V)}{(\Delta V)^2}} \leq \exp\left(\gamma T \theta^2 e^{\theta \Delta V}\right) \leq e^{\theta T}.$$

It suffices to set $\omega_0 = H^* + 1$

From (50) again and the fact that $H \leq -2$ on $M \setminus C$

$$V(X_{t\wedge\tau}^x) \leq V(x) - 2(t \wedge \tau) + M_{t\wedge\tau}^V(x).$$

Thus

$$e^{\theta(V(X_{t\wedge\tau}^x) + t\wedge\tau)} \leq e^{\theta V(x)} Z_{t\wedge\tau}(\theta) e^{(t\wedge\tau)(\gamma\theta^2 e^{\theta\Delta V} - \theta)} \leq e^{\theta V(x)} Z_{t\wedge\tau}(\theta)$$

where $(Z_t(\theta))$ is the supermartingale given by (56). This proves the second assertion. $\square$

Lemma 8.5 *Let*

$$M_T^1(x) = \int_0^T (H(X_s^x) - P_s H(x)) ds.$$

For all $\varepsilon > 0$ there exists $c > 0$ such that for all $x \in C, 0 \leq \theta \leq \theta_0$ and $T \geq 1$

$$\mathsf{E}(e^{\theta M_T^1(x)}) \leq e^{\theta T(\varepsilon + c\theta T)}.$$

Proof: follows from the two following claims.

Claim 1: Let $\mathcal{M}$ be a uniformly integrable family of random variables, centered (i.e $\mathsf{E}(M) = 0$ for all $M \in \mathcal{M}$) and bounded from above (i.e $M \leq c_0 < \infty$ for all $M \in \mathcal{M}$). Then for every $\varepsilon > 0$ there exists $c > 0$ such that for all $\theta \geq 0$

$$\mathsf{E}(e^{\theta M}) \leq e^{\theta(\varepsilon + c\theta)}.$$

Proof of Claim 1: Write $\mathsf{E}(e^{\theta M}) = 1 + \mathsf{E}(r(\theta M))$ with $r(u) = e^u - u - 1$. It is easily checked that $0 \leq r(u) \leq -u$ for $u \leq 0$ and $0 \leq r(u) \leq \max(a^2, b^2 e^b)$ for $-a \leq u \leq b, a \geq 0, b \geq 0$. Thus, for all $R > 0$,

$$\mathsf{E}(r(\theta M)) = \mathsf{E}(r(\theta M)\mathbf{1}_{M \leq -R}) + \mathsf{E}(r(\theta M)\mathbf{1}_{M \geq -R})$$

$$\leq -\theta \mathsf{E}(M\mathbf{1}_{M\leq -R}) + \theta^2 \max(R^2, c_0^2 e^{\theta_0 c_0}))$$

By uniform integrability choose R large enough so that $\mathsf{E}(|M|\mathbf{1}_{M\leq -R}) \leq \varepsilon$ and set $c = \max(R^2, c_0^2 e^{\theta_0 c_0}))$. Then $\mathsf{E}(e^{\theta M}) \leq 1 + \varepsilon\theta + c\theta^2 \leq e^{\varepsilon\theta + c\theta^2}$.

Claim 2: The family $\mathcal{M} = \{\frac{M_T^1(x)}{T} : x \in C, T \geq 1\}$ is uniformly integrable, centered and bounded from above.

Proof of Claim 2: Set, for $p \geq 1$, $\|H\|_{T,p}(x) := (\frac{1}{T}\int_0^T P_s|H|^p(x)ds)^{1/p}$. Then,

$$\frac{M_1^T(x)}{T} \leq H^* + \|H\|_{T,1}(x),\ (\mathsf{E}(|\frac{M_T^1(x)}{T}|^q))^{1/q} \leq \|H\|_{T,q}(x) + \|H\|_{T,1}(x)$$

and, by Hypothesis 4 (ii)' and Theorem 2.1,

$$\|H\|_{T,1}(x) \leq \|H\|_{T,q}(x) \leq [cst(1 + W(x)/T)]^{1/q}.$$

This latter quantity being bounded for $x \in C, T \geq 1$ this proves the claim.

□

Lemma 8.6 *For every $T_0 > 0$ (sufficiently large) and $T_1 > T_0$, there exist positive numbers $\theta \leq \theta_0, \delta, \kappa$ and $\rho < 1$ such that for all $T \in [T_0, T_1]$*

$$P_T(e^{\theta V})(x) \leq \begin{cases} \rho e^{\theta V(x)} \text{ on } M_0^\delta \cap C, \\ \\ \kappa \text{ on } (M_+ \cap C) \setminus M_0^\delta \end{cases}$$

Proof: By compactness of $M_0 \cap C$, the first assertion of Proposition 8.1 remains valid if M_0^δ is replaced by $M_0^\delta \cap C$. The proof of the Lemma is then similar to the proof of Proposition 8.2. It suffices to replace inequality (55) by the inequality given in Lemma 8.5 and to set $\kappa = \sup_{x\in C\cap M_+\setminus M_0^\delta} e^{\theta V(x)} e^{\theta\omega_0 T_1}$ with ω_0 given by Lemma 8.4. □

We now prove Proposition 8.3. Relying on Lemma 8.6 fix T_0, T_1 such that

$$T_1 \geq (2\omega_0 + 1)T_0$$

where ω_0 is given by Lemma 8.4, and let θ, ρ, κ be given by Lemma 8.6.

Let $T \in [\frac{T_0+T_1}{2}, T_1]$. Using the strong Markov property, Lemma 8.6 implies

$$\mathsf{E}_x(e^{\theta V(X_T)}|\mathcal{F}_\tau) \leq \rho e^{\theta V(X_\tau)} + \kappa \leq \rho e^{\theta(V(X_\tau)+\tau)} + \kappa$$

on the event $\tau \leq T - T_0$; and Lemma 8.4 (i) implies that

$$E_x(e^{\theta V(X_T)}|\mathcal{F}_\tau) \leq e^{\theta V(X_\tau)} e^{\omega_0(T-\tau)} \leq e^{\theta(V(X_\tau)+\tau)} e^{\theta\omega_0 T_0} e^{-\theta(T-T_0)}$$

on the even $T - T_0 < \tau \leq T$. Thus, by Lemma 8.4 (i)

$$E_x(e^{\theta V(X_T)} \mathbf{1}_{\tau \leq T}) \leq \max(\rho, e^{\theta((\omega_0+1)T_0 - T)}) e^{\theta V(x)} + \kappa.$$

Also, by Lemma 8.4 (ii)

$$E_x(e^{\theta V(X_T)} \mathbf{1}_{\tau > T}) \leq e^{-\theta T} e^{\theta V(x)}.$$

Replacing ρ by $\max\{\rho, e^{-\theta(T_0+T_1)/2}, e^{-\theta(T_1 - T_0(2\omega_0+1))/2}\}$ and T_0 by $(T_0 + T_1)/2$ proves the result.

Proof of Theorem 4.12 The proof of Theorem 4.12 follows from Proposition 8.3. The argument is verbatim the same as in the proof of Theorem 4.10.

9 Appendix

9.1 Proof of Theorem 2.1

The following Lemma is folklore and will be used repeatedly.

Lemma 9.1 *Let W be a nonnegative proper map, $C \geq 0$ and let $(\mu_n) \subset \mathcal{P}(M)$ be such that $\limsup_{n\to\infty} \mu_n W \leq C$. Then,*

(i) *The sequence (μ_n) is tight and every limit point μ of (μ_n) verifies $\mu W \leq C$.*

(ii) *Let $H : M \mapsto \mathbb{R}$ be a continuous function such that $\frac{W}{1+|H|}$ is proper. If $\mu_n \Rightarrow \mu$ then $\mu_n H \to \mu H$.*

Proof: Assertion (i) easily follows from Markov inequality and monotone convergence.

(ii). Let $G = \frac{W}{1+|H|}$. For all $R \in \mathbb{R} \setminus D_G$ with D_G at most countable, $\mu\{G = R\} = 0$ and, therefore,

$$\lim_{n\to\infty} \mu_n(H\mathbf{1}_{G \leq R}) = \mu(H\mathbf{1}_{G \leq R}).$$

On the other hand $\mu_n(|H|\mathbf{1}_{G>R}) \leq \mu_n(\frac{W}{G}\mathbf{1}_{G>R}) \leq \frac{1}{R}\mu_n(W)$. Thus

$$\lim_{R\to\infty} \limsup_{n\to\infty} \mu_n(|H|\mathbf{1}_{G>R}) = 0$$

and, similarly,

$$\lim_{R\to\infty} \mu(|H|\mathbf{1}_{G>R}) = 0.$$

This proves the result □

We now pass to the proof of Theorem 2.1.

(i). Assumption (i) of Hypothesis 3 makes the process

$$M_t = W(X_t^x) - W(x) - \int_0^t LW(X_s^x)ds, t \geq 0$$

a square integrable martingale satisfying the strong law of large numbers: $\lim_{t\to\infty} \frac{M_t}{t} = 0$ a.s. (the proof is verbatim the same as the proof of Lemma 7.4 detailed above). Thus, using condition (ii) of Hypothesis 3,

$$0 \leq W(X_t^x) + \int_0^t \tilde{W}(X_s^x)ds \leq W(x) + Ct + M_t.$$

Taking the expectation and using Tonelli's Theorem proves assertion (i).

(ii). Dividing by t and letting $t \to \infty$ proves that $\limsup_{t\to\infty} \Pi_t^x \tilde{W} \leq C$ P a.s. Tightness follows from Lemma 9.1.

It remains to show that limit points of (Π_t^x) are invariant probabilities. For Feller discrete time Markov chains, this is a classical result (see e.g [22], Proposition 6.1.8). The proof easily adapts to the present setting as follows.

We claim that for each $f \in C_b(M)$ and $r > 0$ there exists a full measure set $\Omega_{f,r} \in \mathcal{F}$ such that for all $\omega \in \Omega_{f,r}$ $\lim_{t\to\infty} \Pi_t^x(\omega)f - \Pi_t^x(\omega)P_r f = 0$.

Assume the claim is proved. Let $\mathcal{S} \subset C_0(M)$ be a countable dense subset of $C_0(M)$ (recall that $C_0(M)$ is separable) and $\Omega' = \bigcap_{f\in\mathcal{S}, r\geq 0, r\in\mathbb{Q}} \Omega_{f,r}$. Then, by density of $\mathcal{S}$, continuity of $r \mapsto P_r f(x)$ (Hypothesis 2) and dominated convergence, $\mu(\omega)P_r f = \mu(\omega)f$ for all $f \in C_0(M), r \geq 0, \omega \in \Omega'$ and $\mu(\omega)$ a limit point of $\{\Pi_t^x(\omega)\}_{t\geq 0}$. This proves the result.

We now prove the claim. Replacing (X_t) by with (X_{tr}) we can always assume that $r = 1$. Set

$$Qf(x) = \int_0^1 P_s f(x)ds,\ U_{k+1} = \int_k^{(k+1)} f(X_s)ds,$$

$$M_n = \sum_{k=0}^{n-1} (U_{k+1} - Qf(X_k)),\ N_n = \sum_{k=0}^{n-1} (Qf(X_{k+1}) - P_1 Qf(X_k)).$$

The sequences (M_n) and (N_n) are martingales with bounded increments with respect to $\{\mathcal{F}_n\}$. Thus, by the strong law of large number for martingales, $\lim_{n\to\infty} \frac{1}{n} M_n = \lim_{n\to\infty} \frac{1}{n} N_n = 0$ P a.s. Thus

$$\lim_{n\to\infty} \Pi_n^x f - \tilde{\Pi}_n^x Qf = \lim_{n\to\infty} \tilde{\Pi}_n^x Qf - \tilde{\Pi}_n^x P_1 Qf = 0$$

P a.s, where $\tilde{\Pi}_n^x = \frac{1}{n}\sum_{k=0}^n \delta_{X_k^x}$. Replacing f by $P_1 f$ also gives

$$\lim_{n\to\infty} \Pi_n^x P_1 f - \tilde{\Pi}_n^x Q P_1 f = 0$$

P a.s. Since $P_1 Qf = QP_1 f$ we then get that

$$\lim_{n\to\infty} \Pi_n^x f - \Pi_n^x P_1 f = 0.$$

P a.s. The claim is proved.

Probability μ is invariant if and only if $\mu P_t f = \mu f$ for all t and $f \in C_b(M)$. Thus, by Feller continuity, $\mathcal{P}_{inv}(M)$ is closed and compactness equates tightness. The latter will follow from Lemma 9.1 once we have proved that $\mu\tilde{W} \leq C$ for all $\mu \in \mathcal{P}_{inv}(M)$. Let $\mu \in \mathcal{P}_{inv}(M)$. First assume μ ergodic. Then, by Birkhoff ergodic Theorem, $\Pi_t^x \Rightarrow \mu$ for μ almost every x and $\mathbb{P}_x$ almost surely. Thus, $\mu\tilde{W} \leq C$ by Lemma 9.1 (i). If now μ is invariant, the ergodic decomposition theorem, implies that $\mu\tilde{W} \leq C$. This concludes the proof of assertion (ii).

(iii). Set $w(t) = P_t W(x)$. Using the semigroup property and Fubini-Tonelli, we get that

$$w(t+s) - w(t) \leq -\alpha \int_t^{t+s} w(r)dr + Cs \leq Cs \tag{59}$$

$$w(t) - w(t-u) \leq -\alpha \int_{t-u}^{t} w(r)dr + Cs \leq Cu \tag{60}$$

for all $t \geq 0, s \geq 0$ and $0 \leq u \leq t$. On the other hand, by Fatou Lemma and right continuity of $t \to W(X_t^x)$

$$\liminf_{s\to 0, s>0} w(t+s) = \liminf_{s\to 0, s>0} \mathsf{E}(W(X_{t+s}^x)) \geq \mathsf{E}(W(X_t^x)) = w(t).$$

Combined with (59) this shows that $t \to w(t)$ is right-continuous. From (60) we also get that $t \to w(t)$ is lower semi continuous. Set $\Delta^+ w(t) = \limsup_{s\to 0, s>0} \frac{w(t+s)-w(t)}{s}$ and $\Delta^- w(t) = \limsup_{s\to 0, s>0} \frac{w(t)-w(t-s)}{s}$. Using (59) and right continuity, we get that

$$\Delta^+ w(t) \leq -\alpha w(t) + C.$$

Using (60) and lower semi continuity we get that

$$\Delta^- w(t) \leq -\alpha w(t) + C.$$

Set now $\tilde{w}(t) = e^{\alpha t}(w(t) - \frac{C}{\alpha}) - \epsilon t$ for some $\varepsilon > 0$. Then, defining $\Delta^{+,-}\tilde{w}$ like $\Delta^{+,-}w$ with $\tilde{w}$ in place of w we get that

$$\Delta^+ \tilde{w}(t) \leq -\varepsilon \text{ and } \Delta^- \tilde{w}(t) \leq -\varepsilon.$$

This implies that for all $t \geq 0$ there exists an open subset of $\mathbb{R}^+$, I_t containing t such that $w(s) \leq w(t)$ for all $s \in I_t$. In particular the set $\{t \geq 0 : \tilde{w}(t) \leq \tilde{w}(0)\}$ is open in $\mathbb{R}^+$. By lower semi continuity of $\tilde{w}$, it also closed. Being nonempty it equals $\mathbb{R}^+$ by connectedness. Thus $\tilde{w}(t) \leq \tilde{w}(0)$ for all t. Since ε is arbitrary this leads to

$$P_t W(x) = w(t) \leq e^{-\alpha t}(w(0) - \frac{C}{\alpha}) + \frac{C}{\alpha}.$$

□

9.2 Proof of Proposition 3.1

(i). By local Lipschitz continuity and classical results on stochastic differential equations, there exists for any $x \in \mathbb{R}^n$ a unique continuous process (X_t^x) defined on some interval $[0, \tau^x[$ solution to (9), with initial condition $X_0^x = x$ and such that $t < \tau^x \Leftrightarrow \|X_t^x\| < \infty$ (see e.g [65] Chapter IX, exercise 2.10). Furthermore, it is easily checked (by Ito formula and uniqueness of the solutions) that

$$X_{t,i} = x_i \exp\left(\int_0^t [X_{s,i}^{\alpha_i - 1} F_i(X_s) - \frac{1}{2} X_{s,i}^{2(\beta_i - 1)} a_{ii}(X_s)] ds + \sum_j \int_0^t X_{s,i}^{\beta_i - 1} \Sigma_i^j(X_s) dB_s^j\right)$$

where, to shorten notation, X_t stands for X_t^x. Thus

$$x_i > 0 \Rightarrow X_{t,i}^x > 0 \text{ for all } t \in [0, \tau^x[\tag{61}$$

and

$$x_i = 0 \Rightarrow X^x_{t,i} = 0 \text{ for all } t \in [0, \tau^x[\tag{62}$$

We shall now prove that $\tau^x = \infty$.

For any C^2 function $\psi : \mathbb{R}^n \mapsto \mathbb{R}$, by Ito formulae,

$$\psi(X^x_t) - \psi(x) - \int_0^t L\psi(X^x_s)ds = \sum_i \int_0^t \frac{\partial \psi}{\partial x_i}(X^x_s) \left[(X^x_{s,i})^{\beta_i} \sum_{j=1}^m \Sigma^j_i(X^x_s) \right] dB^j_s, \tag{63}$$

Let $\tau^x_k = \inf\{t \geq 0 : \ U(X^x_t) \geq k\}$ for all $k \in \mathbb{N}$. By the assumption on U, for all $x \in \mathbb{R}^n_+$,

$$LU(x) \leq -\alpha U(x) + \beta.$$

Thus

$$\begin{aligned}
k\mathsf{P}(\tau^x_k \leq t) &= \mathsf{E}(U(X^x_{\tau^x_k})\mathbf{1}_{\tau^x_k \leq t}) && (64)\\
&\leq \mathsf{E}(U(X^x_{t\wedge\tau^x_k})) = U(x) + \mathsf{E}(\int_0^{t\wedge\tau^x_k} LU(X^x_s)ds) && (65)\\
&\leq U(x) - \alpha\mathsf{E}(\int_0^{t\wedge\tau^x_k} U(X^x_s)ds) + \beta t && (66)\\
&\leq U(x) + \beta t && (67)
\end{aligned}$$

Hence

$$\mathsf{P}(\tau^x \leq t) = \mathsf{P}(\cap_{k\geq 0}\{\tau^x_k \leq t\}) = \lim_{k\to\infty} \mathsf{P}(\tau^x_k \leq t) = 0$$

proving that $\tau^x = \infty$ almost surely.

We now let (P_t) denote the semigroup acting on bounded (respectively non-negative) measurable functions $f : \mathbb{R}^n_+ \mapsto \mathbb{R}$, by $P_tf(x) = \mathsf{E}(f(X^x_t))$. $C_b(M)-$ Feller continuity just follows from Lebesgue dominated convergence theorem and the continuity in x of the solution X^x_t.

(ii). Inequalities (66, 67) and monotone convergence imply that

$$P_tU(x) = \mathsf{E}(U(X^x_t)) = \lim_{k\to\infty} \mathsf{E}(U(X^x_t)\mathbf{1}_{\tau^x_k \geq t}) \leq U(x) - \alpha\mathsf{E}(\int_0^t U(X^x_s)ds) + \beta t$$

$$= U(x) - \alpha \int_0^t P_sU(x)ds + \beta t \leq U(x) + \beta t$$

where the last equality follows from Fubini-Tonelli theorem. Thus, reasoning exactly like in the proof of Theorem 2.1 (iii) we get that

$$P_tU(x) \leq e^{-\alpha t}(U(x) - \beta/\alpha) + \beta/\alpha.$$

(iii). Let $\psi \in C_c^2(M)$. By Ito formulae $\psi(X_t^x) - \psi(x) - \int_0^t L\psi(X_s^x)ds$ is a Martingale. Thus, taking the expectation, $P_t\psi(x) - \psi(x) = \int_0^t P_s(L\psi)(x)ds$. Thus $|P_t(\psi)(x) - \psi(x)| \leq t\|L\psi\|$ and

$$\lim_{t\to 0} \frac{P_t\psi(x) - \psi(x)}{t} = L\psi(x).$$

This proves that $\psi \in \mathcal{D}(\mathcal{L})$ and $L\psi = \mathcal{L}\psi$. Replacing ψ by ψ^2 shows that $\psi \in \mathcal{D}^2(\mathcal{L})$ and $\Gamma(\psi) = \Gamma_L(\psi)$.

(iv) is immediate from (61) and (62).

(v). For any smooth function $h : \mathbb{R}^+ \mapsto \mathbb{R}$

$$L(h(U)) = h'(U)LU + \frac{1}{2}h''(U)\Gamma_L(U).$$

If h is concave and nondecreasing, this gives

$$L(h(U)) \leq h'(U)LU \leq -\alpha h'(U)U(1+\varphi) + \beta h'(1).$$

Set $h(t) = t^{\frac{1-\eta}{2}}$ and $W = h(U)$. Then $h'(t)t = \frac{1-\eta}{2}h(t)$. Thus

$$L(W) \leq \frac{1-\eta}{2}(-\alpha W(1+\varphi) + \beta).$$

Now

$$\Gamma_L(W) = h'(U)^2\Gamma_L(U) = (\frac{1-\eta}{2})^2 U^{-\eta-1}\Gamma_L(U).$$

Thus

$$\Gamma_L(W) \leq cst(1+U).$$

Let $B : \mathbb{R} \mapsto \mathbb{R}$ be a smooth function such that $B(t) = t$ for $t \leq 1$, $B(t) = 2$ for $t \geq 3$, and $0 \leq B'(t) \leq 1$. Set $W_n = nB(W/n)$ Then $W_n \in \mathcal{D}^2(\mathcal{L})$ (since $W_n - 2n \in C_c^2(M)$), $W_n(x) = W(x)$ and $\mathcal{L}W_n(x) = LW(x)$ whenever $W(x) \leq n$. On the other hand $\Gamma(W_n)(x) = B'^2(W/n)\Gamma_L(W)(x) \leq \Gamma_L(W)(x)$. Thus

$$\sup_{t\geq 0,n} P_t(\Gamma(W_n))(x) \leq \sup_{t\geq 0} P_t\Gamma_L(W)(x) \leq cst(1 + \sup_{t\geq 0} P_tU(x)) < \infty$$

where the last inequality follows from (ii).

References

[1] Y. Bakhtin, T. Hurth, S. D. Lawley, and J. C. Mattingly, *Smooth invariant densities for random switching on the torus*, ArXiv e-prints (2017).

[2] Y. Bakhtin, T. Hurth, and J.C. Mattingly, *Regularity of invariant densities for 1d-systems with random switching*, Nonlinearity **28** (2015), 3755–3787.

[3] Yuri Bakhtin and Tobias Hurth, *Invariant densities for dynamical systems with random switching*, Nonlinearity (2012), no. 10, 2937–2952.

[4] M. Benaïm, *Stochastic persistence*, Preprint, 2014.

[5] M. Benaïm, F. Colonius, and R. Lettau, *Supports of Invariant Measures for Piecewise Deterministic Markov Processes*, NonLinearity (2017), no. 9, 3400–3418.

[6] M. Benaïm, J. Hofbauer, and W. Sandholm, *Robust permanence and impermanence for the stochastic replicator dynamics*, Journal of Biological Dynamics **2** (2008), no. 2, 180–195.

[7] M. Benaïm, T. Hurth, and E. Strickler, *A user-friendly condition for exponential ergodicity in randomly switched environments*, ArXiv e-prints (2018).

[8] M. Benaïm, S. Le Borgne, F. Malrieu, and P-A. Zitt, *Quantitative ergodicity for some switched dynamical systems*, Electronic Communications in Probability **17** (2012), no. 56, 1–14.

[9] ———, *Qualitative properties of certain piecewise deterministic markov processes*, Annales de l'IHP **51** (2015), no. 3, 1040 – 1075.

[10] M. Benaïm and C. Lobry, *Lotka Volterra in fluctuating environment or "how switching between beneficial environments can make survival harder"*, Annals of Applied Probability (2016), no. 6, 3754–3785.

[11] M. Benaïm and S. Schreiber, *Persistence of structured populations in random environments*, Theoretical Population Biology **76** (2009), 19–34.

[12] M. Benaïm and E. Strickler, *Random Switching between Vector Fields Having a Common Zero*, ArXiv e-prints (2017).

[13] J. M Bony, *Principe du maximum, inégalité de harnack et unicité du problème de cauchy pour les opérateurs elliptiques dégénérés*, Annales de l'Institut Fourier **19** (1969), no. 1, 277–304.

[14] P. Chesson, *Multispecies competition in variable environments*, Theoretical Population Biology **45** (1994), 227–276.

[15] P. L. Chesson, *Annual review of ecology and systematics*, Predator-prey theory and variability **9** (1978), 323–347.

[16] ______, *The stabilizing effect of a random environment*, Journal of Mathematical Biology **15** (1982), 1–36.

[17] P. L. Chesson, *Mechanisms of maintenance of species diversity*, Annual Review of Ecology and Systematics **31** (2000), 343–366.

[18] P. L. Chesson and S. Ellner, *Invasibility and stochastic boundedness in monotonic competition models*, Journal of Mathematical Biology **27** (1989), 117–138.

[19] P.L. Chesson and R.R. Warner, *Environmental variability promotes coexistence in lottery competitive systems*, The American Naturalist **117** (1981), 923.

[20] B. Cloez and M. Hairer, *Exponential ergodicity for markov processes with random switching*, Bernoulli (2015), no. 1, 505–536.

[21] M. H. A. Davis, *Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models*, J. Roy. Statist. Soc. Ser. B **46** (1984), no. 3, 353–388, With discussion. MR 790622

[22] M. Duflo, *Random Iterative Models*, Springer Paris, 2000.

[23] Richard Durrett, *Stochastic calculus*, Probability and Stochastics Series, CRC Press, Boca Raton, FL, 1996, A practical introduction. MR 1398879

[24] E. B. Dynkin, *Markov processes. Vols. I, II*, Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone. Die Grundlehren der Mathematischen Wissenschaften, Bände 121, vol. 122, Academic Press Inc., Publishers, New York; Springer-Verlag, Berlin-Göttingen-Heidelberg, 1965. MR 0193671

[25] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley Series in Probability and Statistics, 1986.

[26] Steven N. Evans, Alexandru Hening, and Sebastian J. Schreiber, *Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments*, J. Math. Biol. **71** (2015), no. 2, 325–359. MR 3367678

[27] ______, *Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments*, J. Math. Biol. **71** (2015), no. 2, 325–359. MR 3367678

[28] Dean Foster and Peyton Young, *Stochastic evolutionary game dynamics*, Theoret. Population Biol. **38** (1990), no. 2, 219–232. MR 1075000

[29] H. Freedman and P. Waltman, *Mathematical analysis of some three species food-chain models*, Math. Biosci. **33** (1977), 257–276.

[30] D. Fudenberg and C. Harris, *Evolutionary dynamics with aggregate shocks*, J. Econom. Theory **57** (1992), no. 2, 420–441. MR 1180005

[31] B. M. Garay and J. Hofbauer, *Robust permanence for ecological equations, minimax, and discretization*, Siam J. Math. Anal. **Vol. 34** (2003), no. 5, 1007–1039.

[32] Thomas C. Gard, *Persistence in food chains with general interactions*, Math. Biosci. **51** (1980), no. 1-2, 165–174. MR 605583

[33] ______, *Persistence in food webs: Holling-type food chains*, Math. Biosci. **49** (1980), no. 1-2, 61–67. MR 572842

[34] Thomas C. Gard and Thomas G. Hallam, *Persistence in food webs. I. Lotka-Volterra food chains*, Bull. Math. Biol. **41** (1979), no. 6, 877–891. MR 640001

[35] Shengliang Guo and Yijun Hu, *Asymptotic behavior and numerical simulations of a Lotka-Volterra mutualism system with white noises*, Adv. Difference Equ. (2017), Paper No. 125, 19. MR 3645719

[36] M. Hairer and J. Mattingly, *Seminar on stochastic analysis, random fields and applications vi*, Progress in Probability, ch. Yet another look at HarrisŠ ergodic theorem for Markov chains, pp. 109–117, Springer, 2011.

[37] A. Hening and D. H. Nguyen, *Coexistence and extinction for stochastic Kolmogorov systems*, ArXiv e-prints (2017).

[38] A. Hening and E. Strickler, *On a predator-prey system with random switching that never converges to its equilibrium*, ArXiv e-prints (2017).

[39] M. W. Hirsch and H. Smith, *Monotone dynamical systems*, Handbook of Differential Equations (A Cañada, P Drábek, and A Fonda, eds.), vol. 2, ELSEVIER, 2005.

[40] Morris W. Hirsch, *Systems of differential equations which are competitive or cooperative. III. Competing species*, Nonlinearity **1** (1988), no. 1, 51–71. MR 928948

[41] J. Hofbauer, *A general cooperation theorem for hypercycles*, Monatshefte fur Mathematik (1981), no. 91, 233–240.

[42] J. Hofbauer and K. Sigmund, *Evolutionary games and population dynamics*, Cambridge University Press, 1998.

[43] Josef Hofbauer and Lorens A. Imhof, *Time averages, recurrence and transience in the stochastic replicator dynamics*, Ann. Appl. Probab. **19** (2009), no. 4, 1347–1368. MR 2538073

[44] Josef Hofbauer and Sebastian J. Schreiber, *To persist or not to persist?*, Nonlinearity **17** (2004), no. 4, 1393–1406. MR 2069711

[45] G. E. Hutchinson, *The paradox of the plankton*, The American Naturalist **95** (1961), no. 882, 137–145.

[46] Kanji Ichihara and Hiroshi Kunita, *A classification of the second order degenerate elliptic operators and its probabilistic characterization*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **30** (1974), no. 3, 235–254.

[47] Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1981. MR 637061

[48] O. Kallenberg, *Foundations of Modern Probability Second edition*, Probability and its Applications, Springer, 2001.

[49] Russell Lande, Steinar Engen, and Bernt-Erik Saether, *Stochastic population dynamics in ecology and conservation*, Oxford University Press on Demand, 2003.

[50] J. F. Le Gall, *Mouvement brownien, martingales et calcul stochastique*, Springer, 2013.

[51] ______, *Brownian motion, martingales, and stochastic calculus*, Springer, 2016.

[52] Dan Li, Shengqiang Liu, and Jing'an Cui, *Threshold dynamics and ergodicity of an SIRS epidemic model with markovian switching*, Journal of Differential Equations **263** (2017), no. 12, 8873–8915.

[53] G. SH. Lipster, *A strong law of large numbers for local martingales*, Stochastics **3** (1980), 217–228.

[54] F. Malrieu, *Some simple but challenging Markov processes*, Ann. Fac. Sci. Toulouse Math. (6) **24** (2015), no. 4, 857–883. MR 3434260

[55] F. Malrieu and P.-A. Zitt, *On the persistence regime for Lotka-Volterra in randomly fluctuating environments*, ArXiv e-prints (2016).

[56] Robert M. May and Warren J. Leonard, *Nonlinear aspects of competition between three species*, SIAM J. Appl. Math. **29** (1975), no. 2, 243–253, Special issue on mathematics and the social and biological sciences. MR 0392035

[57] S.P̃. Meyn and R.L̃. Tweedie, *Markov Chains and Stochastic Stability, Second Edition.*, Cambridge University Press, 2009.

[58] Janusz Mierczyński, *The C^1 property of carrying simplices for a class of competitive systems of ODEs*, J. Differential Equations **111** (1994), no. 2, 385–409. MR 1284419

[59] ______, *On smoothness of carrying simplices*, Proc. Amer. Math. Soc. **127** (1999), no. 2, 543–551. MR 1606000

[60] J. Mierczyski, W. Shen, and X.Q. Zhao, *Uniform persistence for nonautonomous and random parabolic kolmogorov systems*, Journal of Differential Equations (2004), no. 204, 471–510.

[61] S. Méléard and D. Villemonais, *Quasi-stationary distributions and population processes*, Probability Surveys **9** (2012), 340–410.

[62] K. Sigmund P. Schuster and R. Wolff, *Dynamical systems under constant organization 3: Cooperative and competitive behavior of hypercycles*, Journal of Differential Equations **32** (1979), 357–368.

[63] G. Da Prato and M. Röckner, *Cores for generators of some markov semigroups*, fix **fix**, fix.

[64] E. Priola, *On a class of markov type semigroups in spaces of uniformly continuous and bounded functions*, Studia Math **136** (1999), 271–295.

[65] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer Verlag, 2004.

[66] M. Rosenzweig and R. MacArthur, *Graphical representation and stability con- ditions of predator-prey interaction*, American Naturalist **97** (1963), 209–223.

[67] G. Roth and S.J. Schreiber, *Persistence in fluctuating environments for interacting structured populations*, Journal of Mathematical Biology **68** (2014), 1267–1317.

[68] S. Schreiber, *Criteria for c^r robust permanence*, J Differential Equations (2000), 400–426.

[69] ______, *Persistence for stochastic difference equations: A mini review*, Journal of Difference Equations and Applications **18** (2012), 1381–1403.

[70] S. Schreiber, M. Benaïm, and KAS Atchadé, *Persistence in fluctuating environments*, Journal of Mathematical Biology **62** (2011), 655–683.

[71] Sebastian J Schreiber, *Coexistence in the face of uncertainty*, Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science, Springer, 2017, pp. 349–384.

[72] H. L. Smith, *The rosenzweig-macarthur predator-prey model*, https://math.la.asu.edu/ halsmith/Rosenzweig.pdf.

[73] H. L. Smith and H. R. Thieme, *Dynamical systems and population persistence*, vol. 118, American Mathematical Society, Providence, RI, 2011.

[74] D. W. Stroock and S. R. S Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Berkeley, 1970-1971) publisher= Berkeley. University of California Press., vol. 3, 1972, pp. 333–359.

[75] H.R. Thieme, *Uniform persistence and permanence for non-autonomous semi ows in population biology*, Mathematical Biosciences (2000), no. 166, 173–201.

[76] Michael Turelli, *Random environments and stochastic calculus*, Theoretical population biology **12** (1977), no. 2, 140–178.

[77] E. C. Zeeman and M. L. Zeeman, *An n-dimensional competitive Lotka-Volterra system is generically determined by the edges of its carrying simplex*, Nonlinearity **15** (2002), no. 6, 2019–2032. MR 1938479

[78] Xiao-Qiang Zhao, Jonathan Borwein, and Peter Borwein, *Dynamical systems in population biology*, Springer, 2017.

Acknowledgments

This work is supported by the SNF grant 200021_15728. I thank Jean Baptiste Bardet, Patrick Cattiaux, Alex Hening, Tobias Hurt, Eva Locherbach, Florent Malrieu, Janusz Mierczynski, Sebastian Schreiber, Edouard Strickler, Pierre André Zitt for valuable discussions on different topics related to this paper. Special thanks to Edouard Strickler for his help with the simulation of the May Leonard process with Scilab and to Jean Baptiste Bardet for his help with the computations of Lie Brackets using the formal software Python/Sympy.