Random Switching between Vector Fields Having a Common Zero

Michel Benaïm, Edouard Strickler
Institut de Mathématiques
Université de Neuchâtel, Switzerland

February 10, 2017

Abstract

Let E be a finite set, $\{F^i\}_{i \in E}$ a family of vector fields on $\mathbb{R}^d$ leaving positively invariant a compact set M and having a common zero $p \in M$. We consider a piecewise deterministic Markov process (X, I) on $M \times E$ defined by $\dot{X}_t = F^{I_t}(X_t)$ where I is a jump process controlled by $X : \mathsf{P}(I_{t+s} = j | (X_u, I_u)_{u \le t}) = a_{ij}(X_t)s + o(s)$ for $i \neq j$ on $\{I_t = i\}$.

We show that the behavior of (X, I) is mainly determined by the behavior of the linearized process (Y, J) where $\dot{Y}_t = A^{J_t} Y_t$, A^i is the Jacobian matrix of F^i at p and J is the jump process with rates $(a_{ij}(p))$. We introduce two quantities Λ^- and Λ^+ respectively defined as the *minimal* (respectively *maximal*) *growth rate* of $\|Y_t\|$, where the minimum (respectively maximum) is taken over all the ergodic measures of the angular process (Θ, J) with $\Theta_t = \frac{Y_t}{\|Y_t\|}$. It is shown that Λ^+ coincides with the top Lyapunov exponent (in the sense of ergodic theory) of (Y, J) and that under general assumptions $\Lambda^- = \Lambda^+$. We then prove that, under certain irreducibility conditions, $X_t \to p$ exponentially fast when $\Lambda^+ < 0$ and (X, I) converges in distribution at an exponential rate toward a (unique) invariant measure supported by $M \setminus \{p\} \times E$ when $\Lambda^- > 0$. Some applications to certain epidemic models in a fluctuating environment are discussed and illustrate our results.

Keywords: Piecewise deterministic Markov processes; Random Switching; Lyapunov Exponents; Stochastic Persistence; Hypoellipticity, Hörmander-Bracket conditions; Epidemic models; SIS

AMS subject classifications 60J25, 34A37, 37H15, 37A50, 92D30

Contents

1 Introduction

Let E be a finite set and $\mathrm{F} = \{F^i\}_{i \in E}$ a family of C^2 globally integrable vector fields on $\mathbb{R}^d$. For each $i \in E$ we let $\Psi^i = \{\Psi^i_t\}$ denote the flow induced by F^i. We assume throughout that there exists a closed set $M \subset \mathbb{R}^d$ which is *positively invariant* under each Ψ^i. That is

$$\Psi^i_t(M) \subset M$$

for all $t \geq 0$.

Consider a Markov process $Z = (Z_t)_{t \geq 0}, Z_t = (X_t, I_t)$, living on $M \times E$ whose infinitesimal generator acts on functions $g : M \times E \mapsto \mathbb{R}$, smooth in the first variable, according to the formula

$$\mathcal{L}g(x,i) = \langle F^i(x), \nabla g^i(x) \rangle + \sum_{j \in E} a_{ij}(x)(g^j(x) - g^i(x)), \tag{1}$$

where $g^i(x)$ stands for $g(x,i)$ and $a(x) = (a_{ij}(x))_{i,j \in E}$ is an irreducible *rate* matrix continuous in x. Here, by a rate matrix, we mean a matrix having nonnegative off diagonal entries and zero diagonal entries.

In other words, the dynamics of X is given by an ordinary differential equation

$$\frac{dX_t}{dt} = F^{I_t}(X_t), \tag{2}$$

while I is a continuous time jump process taking values in E controlled by X :

$$\mathsf{P}(I_{t+s} = j | \mathcal{F}_t, I_t = i) = a_{ij}(X_t)s + o(s) \text{ for } j \neq i \text{ on } \{I_t = i\},$$

where $\mathcal{F}_t = \sigma((X_s, I_s)\ : s \leq t\}$.

This class of processes belongs to the wider class of *Piecewise Deterministic Markov Processes* (PDMPs), a term coined by Davis [22], and has recently been the focus of much attention. Criteria, based on irreducibility and Hörmander type conditions, ensuring uniqueness and absolute continuity of an invariant probability measure have been obtained by Bakhtin and Hurth [5] for constant jump rates ($a_{ij}(x) = a_{ij}$) and by Benaïm, Le Borgne, Malrieu and Zitt [15] for more general rates. Exponential convergence (in total variation) toward this measure and a support theorem, describing the support of the law of $(Z_t)_{z \geq 0}$ are also proved in [15] when M is compact. In the one dimensional case (i.e $d = 1$) smoothness properties of the invariant measure are thoroughly investigated by Bakhtin, Hurth and Mattingly [6]. When irreducibility fails to hold, the support of invariant probabilities can be determined in terms of invariant control sets of an associated deterministic control system (see Benaïm, Colonius and Lettau [10]). When the vector fields are exponentially asymptotically stable in "average", exponential convergence toward an invariant measure are obtained for Wassertein distances by Benaïm, Le Borgne, Malrieu and Zitt [13], Cloez and Hairer [20]. Several examples, either linear (Benaïm, Le Borgne, Malrieu and Zitt [14], Lawley, Mattingly and Reed [34], Lagasquie [32]), or nonlinear (Benaïm and Lobry [16], Malrieu and Hoa Phu [36]) show that the behavior of the process is not solely determined by the dynamics of the Ψ^i but can be highly sensitive to the switching rates. We refer the reader to the recent overview by Malrieu [35], describing these results among others.

In the present paper we will investigate the behavior of the process Z under the following two conditions:

C1 The origin lies in M and is a common equilibrium:

$$F^i(0) = 0 \text{ for all } i \in E;$$

C2 The set M is compact and *locally star shaped at the origin*, meaning that there exists $\delta > 0$ such that

$$x \in M \text{ and } \|x\| \leq \delta \Rightarrow [0, x] \subset M.$$

Compactness of M is assumed here for simplicity, but some of the results can be extended to noncompact sets provided we can control the behavior of the process near infinity, for instance with a suitable Lyapunov function.

Briefly put, our main result is that the long term behavior of the process is determined by the behavior of the process obtained by linearization at the origin and, under suitable irreducibility and hypoellipticity conditions, by the top Lyapunov exponent of the linearized system. If negative, then $X = (X_t)$ converges almost surely and exponentially fast to zero. If positive, and $X_0 \neq 0$, the empirical occupation measure (respectively the law) of Z converge almost surely (respectively in total variation at an exponential rate) toward a unique probability measure putting zero mass on $\{0\} \times E$. Such a correspondence between the sign of the top Lyapunov exponent and the behavior of nonlinear system is reminiscent of the results obtained by Baxendale [7] and others for Stratanovich stochastic differential equations (see [7] and the references therein, and Hening, Nguyen and Yin [29] for similar recent results in the context of population dynamics).

Our proofs rely, on one hand, on the qualitative theory of PDMPs (as developed in [5] and [15]) and, on the other hand, on some recent results on *stochastic persistence* (Benaïm [9]) strongly inspired by the seminal works of Schreiber, Hofbauer and their co-authors on *persistence*, first developed for purely deterministic systems (Schreiber [40], Garay and Hofbauer [25], Hofbauer and Schreiber [31]) and later for certain stochastic systems (Benaïm, Hofbauer and Sandholm [12], Benaïm and Schreiber [17], Schreiber, Benaïm and Atchade [42], Schreiber [41], Roth and Schreiber [39]).

Our original motivation was to analyze the behavior of certain epidemic models evolving in a fluctuating environment. A famous, and now classical, deterministic model of infection is given by the Lajmanovich and Yorke differential equation ([33]). This equation leaves positively invariant the unit cube of $\mathbb{R}^d$ and models the evolution of the infection level between d groups. Depending on the parameters of the model (the environment), either the disease dies out (i.e all the trajectories converge to the origin) or stabilizes (i.e all non zero trajectories converge toward a unique positive equilibrium). Deterministic switching between several environment have been recently considered by Ait Rami, Bokharaie, Mason and Wirth [1]. The results here allow to describe the behavior of the process when switching between environment evolves randomly. In particular we can produce paradoxical examples for which, although each deterministic dynamics leads to the extinction (respectively persistence) of the disease, the random switching process leads to persistence (respectively extinction) of the disease.

1.1 Outline of contents

Section 2 considers the linearized system (Y, J) where $\dot{Y}_t = A^{J_t} Y_t$, $A^i = DF^i(0)$ (the Jacobian of F^i at 0) and J is the jump process with rate matrix $(a_{ij}) = (a_{ij}(0))$. We introduce two quantities Λ^- and Λ^+ respectively defined as the *minimal* (respectively *maximal*) *growth rate* of $\|Y_t\|$, where the minimum (respectively maximum) is taken over all the ergodic measures of the angular Markov process (Θ, J) with $\Theta_t = \frac{Y_t}{\|Y_t\|}$. It is shown (Proposition 2.4) that Λ^+ coincides with the top Lyapunov exponent (in the sense of ergodic theory) of (Y, J) and some conditions are given ensuring that $\Lambda^- = \Lambda^+$, first for arbitrary A^is (Proposition 2.9) and then for Metzler matrices (Proposition 2.11).

The main results of the paper are stated in Section 3.

- If $\Lambda^+ < 0$, $X_t \to 0$ exponentially fast, locally (i.e for $\|X_0\|$ small enough), with positive probability. If furthermore 0 is accessible, convergence is global and almost sure (Theorem 3.1).
- If $\Lambda^- > 0$ and $X_0 \neq 0$, the process is *persistent* in the sense that weak limit points of its empirical occupation measure are almost surely invariant probabilities over $M \setminus \{0\} \times E$ (Theorem 3.2). If in addition the F^is satisfy a certain Hörmander-type bracket condition at some accessible point, then there is a unique invariant probability on $M \setminus \{0\} \times E$ toward which the empirical occupation measure converges almost surely (Theorem 3.3). Under a strengthening of the bracket condition, the distribution of the process converges also exponentially fast in total variation (Theorem 3.4).

Section 4 discusses some applications of our results to certain epidemic models in a fluctuating environment. The focus is on the situation where the F^is are given by Lajmanovich and Yorke type vector fields [33] (or more generally sub homogeneous cooperative systems in the sense of Hirsch [30]). Several examples are analyzed and a theorem proving exponential convergence of the distribution (for a certain Wasserstein distance) in absence of the bracket condition is stated (Theorem 4.11).

Sections 5 and 6 are devoted to the proofs of Theorems 3.1, 3.2, 3.3, 3.4 and 4.11. The proofs of certain results stated in Section 2 are given in appendix (Section 7) for convenience.

1.2 Notation

The following notation will be used throughout: $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product in $\mathbb{R}^d$, $\| \cdot \|$ the associated norm, $B(x, r) = \{y \in \mathbb{R}^d : \|y - x\| \leq r\}$ the closed ball centered at x with radius r and $S^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$ the unit sphere.

Notation for Markov processes For any polish space $\mathcal{X}$ such as M, S^{d-1}, E, $M \times E$, equipped with its Borel sigma-field, we let $\mathcal{P}(\mathcal{X})$ denote the set of (Borel) probabilities over $\mathcal{X}$. We shall consider below certain Markov processes $\tilde{Z}$ (like Z) taking values in $\mathcal{X}$ with *cadlag* (right continuous, left limit) paths. Given such a process and $\mu \in \mathcal{P}(\mathcal{X})$ we let $\mathbb{P}_\mu^{\tilde{Z}}$ denote the law of $\tilde{Z}$ on the Skorokhod space $D(\mathbb{R}^+, \mathcal{X})$ when $\tilde{Z}_0$ has law μ. As usual, $\mathbb{P}_z^{\tilde{Z}}$ stands for $\mathbb{P}_{\delta_z}^{\tilde{Z}}$ for all $z \in \mathcal{X}$. The *Markov semi-group* induced by $\tilde{Z}$, denoted $(P_t^{\tilde{Z}})_{t \geq 0}$, acts on bounded measurable functions $f : \mathcal{X} \mapsto \mathbb{R}$ according to the formula

$$P_t^{\tilde{Z}} f(z) = \mathbb{E}_z(f(\tilde{Z}_t)) = \int f(\eta(t)) d\mathbb{P}_z^{\tilde{Z}}(\eta).$$

By duality it acts on $\mathcal{P}(\mathcal{X})$ by

$$(\mu P_t^{\tilde{Z}}) f = \mu(P_t^{\tilde{Z}} f),$$

where here and throughout μf stands for $\int f d\mu$. Probability $\mu \in \mathcal{P}(\mathcal{X})$ is said *invariant* for $\tilde{Z}$ provided $\mu P_t^{\tilde{Z}} = \mu$ for all $t \geq 0$. It is called *ergodic* if, in addition of being invariant, the only bounded measurable functions $f : \mathcal{X} \mapsto \mathbb{R}$ for which $\sup_{t \geq 0} \mu(|P_t^{\tilde{Z}} f - f|) = 0$ are μ-almost surely constant.

We let $\mathcal{P}_{inv}^{\tilde{Z}} \subset \mathcal{P}(\mathcal{X})$ denote the (possibly empty) set of *invariant probabilities* of $\tilde{Z}$ and $\mathcal{P}_{erg}^{\tilde{Z}} \subset \mathcal{P}_{inv}^{\tilde{Z}}$ the subset of *ergodic probabilities*. Recall that $\mathcal{P}_{erg}^{\tilde{Z}}$ can also be defined as the set of extremal points of $\mathcal{P}_{inv}^{\tilde{Z}}$.

A key property, that will be used later without further notice, is that whenever $\mu \in \mathcal{P}_{inv}^{\tilde{Z}}$ (respectively $\mu \in \mathcal{P}_{erg}^{\tilde{Z}}$), $\mathbb{P}_\mu^{\tilde{Z}}$ is invariant (respectively ergodic), in the sense of ergodic theory, for the shift $\mathbf{\Theta} = (\mathbf{\Theta}_t)_{t \geq 0}$ on $D(\mathbb{R}^+, \mathcal{X})$; where

$$\mathbf{\Theta}_t(\eta)(s) = \eta(t+s).$$

We refer the reader to Meyn and Tweedie ([37], chapter 17) for a proof and more details.

Accessibility Let $\tilde{\mathrm{F}} = \{\tilde{F}^i\}_{i \in E}$ be a family of bounded C^1 vector fields on $\mathbb{R}^d$ indexed by E. For instance $\tilde{\mathrm{F}} = \mathrm{F}$. We let $\mathsf{co}(\tilde{\mathrm{F}})$ denote the compact convex set valued mapping defined by

$$\mathsf{co}(\tilde{\mathrm{F}})(x) = \{\sum_{j \in E} \alpha_j \tilde{F}^j(x) \ : \alpha_j \geq 0, \sum_{j \in E} \alpha_j = 1\}.$$

Given a closed set $A \subset \mathbb{R}^d$ and $B \subset \mathbb{R}^d$ we say that A is $\tilde{F}$*-accessible* from B if for every neighborhood U of A and every $x \in B$, there exists a (absolutely continuous) function $\eta : \mathbb{R}_+ \mapsto \mathbb{R}^d$, solution to the differential inclusion

$$\begin{cases} \dot{\eta}(t) \subset \mathsf{co}(\tilde{\mathrm{F}})(\eta(t)) \\ \eta(0) = x \end{cases}$$

such that $\eta(t) \in U$ for some $t > 0$. An equivalent formulation (see e.g Theorem 2.2 in [10]) is that A is *reachable* from B by the control system

$$\begin{cases} \dot{x} = \sum_j \tilde{F}^j(x) v_j(t) \\ x(0) = x \end{cases}$$

where the control $v \in D(\mathbb{R}_+, \{e_j\}_{j \in E})$ with $\{e_j\}_{j \in E}$ the canonical basis of $\mathbb{R}^E$. Note that this notion is what is called D-approachability in [5].

2 The Linearized system

Let, for $i \in E, A^i = DF^i(0)$ denote the Jacobian matrix of F^i at the origin. We let $C_M \subset \mathbb{R}^d$ denote the cone defined as

$$C_M = \{tx : \ t \geq 0, x \in M, \|x\| \leq \delta\}$$

where δ is like in condition $C2$.

Lemma 2.1 *For all* $t \geq 0$ $e^{tA^i} C_M \subset C_M$.

Proof Let $x \in C_M$. For ε small enough, by definition of C_M and continuity of Ψ_t^i at 0 $\Psi_t^i(\varepsilon x) \in M \cap B(0,\delta)$. Hence $\frac{\Psi_t^i(\varepsilon x)}{\varepsilon} \in C_M$ and letting $\varepsilon \to 0$ this shows that $D\Psi_t^i(0)x = e^{tA^i}x \in C_M$. **QED**

Define the *linearized system of Z at the origin* as the "linear" PDMP (Y,J) living on $C_M \times E$ whose generator L is given by

$$Lg(y,i) = \langle A^i y, \nabla g^i(y)\rangle + \sum_{j\in E} a_{ij}(g^j(y) - g^i(y)),$$

where

$$a_{ij} = a_{ij}(0).$$

A trajectory $(Y_t, J_t)_{t\geq 0}$ with initial condition (y,i) is then obtained as a solution to

$$\begin{cases} \frac{dY_t}{dt} = A^{J_t} Y_t \\ Y_0 = y, \end{cases} \tag{3}$$

where (J_t) is a continuous time Markov process on E with jump rates (a_{ij}) based at $J_0 = i$.

By irreducibility of (a_{ij}), J has a unique invariant probability $p = (p_i)_{i\in E}$, characterized by

$$\forall i \in E, \sum_j (p_j a_{ji} - p_i a_{ij}) = 0.$$

Whenever $y \neq 0$ the *polar decomposition*

$$(\Theta_t = \frac{Y_t}{\|Y_t\|}, \rho_t = \|Y_t\|) \in S^{d-1} \cap C_M \times \mathbb{R}_+$$

is well defined and (3) can be rewritten as

$$\begin{cases} \frac{d\Theta_t}{dt} = G^{J_t}(\Theta_t) \\ \frac{d\rho_t}{dt} = \langle A^{J_t}\Theta_t, \Theta_t\rangle \rho_t, \end{cases} \tag{4}$$

where for all $i \in E$ G^i is the vector field on S^{d-1} defined by

$$G^i(\theta) = A^i\theta - \langle A^i\theta, \theta\rangle\theta. \tag{5}$$

Remark 2.2 For stochastic differential equations, the idea of introducing, this polar decomposition goes back to Hasminskii [28] and has proved to be a fundamental tool for analyzing linear stochastic differential equations (see e.g [7]), linear random dynamical systems (see e.g chapter 6 of Arnold [2]) and more recently certain linear PDMPs in [14], [34] or [32].

□

With obvious notation, the processes

$$(\Theta, \rho, J) = ((\Theta_t, \rho_t, J_t))$$

and

$$(\Theta, J) = ((\Theta_t, J_t))$$

are two PDMPs respectively living on $S^{d-1} \cap C_M \times \mathbb{R}_+ \times E$ and $S^{d-1} \cap C_M \times E$.

By compactness of $S^{d-1} \cap C_M$ and Feller continuity of (Θ, J) (see [15], Proposition 2.1), $\mathcal{P}_{inv}^{(\Theta,J)}$ is a nonempty compact (for the topology of weak* convergence) subset of $\mathcal{P}(S^{d-1} \cap C_M \times E)$.

2.1 Average growth rates

Define, for each $\mu \in \mathcal{P}_{inv}^{(\Theta,J)}$, the *$\mu$-average growth rate* as

$$\Lambda(\mu) = \int \langle A^i\theta, \theta\rangle \mu(d\theta di) = \sum_{i\in E} \int_{S^{d-1}\cap C_M} \langle A^i\theta, \theta\rangle \mu^i(d\theta), \tag{6}$$

where $\mu^i(.)$ is the measure on $S^{d-1} \cap C_M$ defined by

$$\mu^i(A) = \mu(A \times \{i\}).$$

Note that when μ is ergodic, by equation (4) and Birkhoff ergodic theorem

$$\lim_{t\to\infty} \frac{\log(\rho_t)}{t} = \Lambda(\mu)$$

$\mathbb{P}_\mu^{(\Theta,J)}$ almost surely.

Define similarly the *extremal average growth rates* as the numbers

$$\Lambda^- = \inf\{\Lambda(\mu) \ : \mu \in \mathcal{P}_{erg}^{(\Theta,J)}\} \text{ and } \Lambda^+ = \sup\{\Lambda(\mu) \ : \mu \in \mathcal{P}_{erg}^{(\Theta,J)}\}. \tag{7}$$

The following rough estimate is a direct consequence of (6). Recall that $p = (p_i)_{i\in E}$ is the invariant probability of J.

Lemma 2.3

$$\sum_i p_i\lambda_{min}(\frac{A^i + (A^i)^T}{2}) \leq \Lambda^- \leq \Lambda^+ \leq \sum_i p_i\lambda_{max}(\frac{A^i + (A^i)^T}{2}),$$

where λ_{min} (respectively λ_{max}) denotes the smallest (respectively largest) eigenvalue.

The signs of Λ^- and Λ^+ will play a crucial role for determining the asymptotic behavior of the non linear process Z. But before stating our main results, it is interesting to compare them with the usual Lyapunov exponents given by the multiplicative ergodic theorem.

2.2 Relation with Lyapunov exponents

Set $\Omega = D(\mathbb{R}_+, E)$ and for $\omega \in \Omega$ and $y \in \mathbb{R}^d$, let

$$t \mapsto \varphi(t, \omega)y$$

denote the solution to the linear differential equation

$$\dot{y} = A^{\omega_t}y$$

with initial condition $\varphi(0, \omega)y = y$.

Then, φ is a *linear random dynamical system* over the ergodic dynamical system $(\Omega, \mathbb{P}_p^J, \mathbf{\Theta})$, for which the assumptions of the multiplicative ergodic theorem are easily seen to be satisfied (see e.g [2], Theorem 3.4.1 or Colonius and Mazanti [21]). Thus, according to this theorem, there exist $1 \leq \tilde{d} \leq d$, numbers

$$\lambda_{\tilde{d}} < \ldots < \lambda_1,$$

called *the Lyapunov exponents* of $(\varphi, \mathbf{\Theta})$, a Borel set $\tilde{\Omega} \subset \Omega$ with $\mathbb{P}^J_p(\tilde{\Omega}) = 1$, and for each $\omega \in \tilde{\Omega}$ distinct vector spaces

$$\{0\} = V_{\tilde{d}+1}(\omega) \subset V_{\tilde{d}}(\omega) \subset \ldots \subset V_i(\omega) \ldots \subset V_1(\omega) = \mathbb{R}^d$$

(measurable in ω) such that

$$\lim_{t\to\infty} \frac{1}{t} \log \|\varphi(t,\omega)y\| = \lambda_i \tag{8}$$

for all $y \in V_i(\omega) \setminus V_{i+1}(\omega)$.

Proposition 2.4 *For all $\mu \in \mathcal{P}^{(\Theta,J)}_{erg}$*

$$\Lambda(\mu) \in \{\lambda_{\tilde{d}}, \ldots, \lambda_1\}$$

and

$$\Lambda^+ = \lambda_1.$$

Proof Let $\mu \in \mathcal{P}^{(\Theta,J)}_{erg}$. Then, $\mathbb{P}^{(\Theta,J)}_\mu$ almost surely

$$\lim_{t\to\infty} \frac{1}{t} \log(\|\varphi(t,J)\Theta_0)\|) = \lim_{t\to\infty} \frac{1}{t} \int_0^t \langle A^{J_s}\Theta_s, \Theta_s\rangle ds = \Lambda(\mu)$$

The first equality follows from (3), (4) and the definition of $\varphi(t,\omega)$. The second follows from Birkhoff ergodic theorem. Therefore, there exists a Borel set $\mathcal{B} \subset (S^{d-1} \cap C_M) \times \Omega$ such that for all $(\theta,\omega) \in \mathcal{B}$

$$\lim_{t\to\infty} \frac{1}{t} \log(\|\varphi(t,\omega)\theta\|) = \Lambda(\mu) \tag{9}$$

and $\mathbb{P}^{(\Theta_0,J)}_\mu(\mathcal{B}) = 1$, where $\mathbb{P}^{(\Theta_0,J)}_\mu(d\theta d\omega) = \sum_{i\in E} \mathbb{P}^J_i(d\omega)\mu^i(d\theta)$ is the law of (Θ_0, J) under $\mathbb{P}^{(\Theta,J)}_\mu$.

Let $\tilde{\Omega} \subset \Omega$ be the set given by the multiplicative ergodic theorem and $\tilde{\mathcal{B}} = \{(\theta,\omega) \in \mathcal{B} : \omega \in \tilde{\Omega}\}$. Then $\mathbb{P}^{(\Theta_0,J)}_\mu(S^{d-1} \cap C_M \times \tilde{\Omega}) = \mathbb{P}^J_\mu(\tilde{\Omega}) = 1$. Hence $\mathbb{P}^{(\Theta_0,J)}_\mu(\tilde{\mathcal{B}}) = 1$ and for all $(\theta,\omega) \in \tilde{\mathcal{B}}$ the left hand side of equality (9) equals λ_i for some i.

It remains to show that $\lambda_1 = \Lambda^+$. Replace Ω by $\Omega' = D(\mathbb{R}, E)$ and consider the two-sided random dynamical system $\varphi(t,\cdot)$ over $(\Omega', \mathbb{P}^J_p, \mathbf{\Theta})$. By the two-sided version of the multiplicative ergodic theorem (see [2], Theorem 3.4.11 and Remark 3.4.14), there exists a set $\tilde{\Omega} \subset \Omega'$ with $\mathbb{P}^J_p(\tilde{\Omega}) = 1$, and for each $\omega \in \tilde{\Omega}$ a random vector space $E_1(\omega)$ measurable with respect to the σ field $\mathcal{F}_0 = \sigma\{\omega_t : t \leq 0\}$ such that, when $i = 1$, equation (8) holds for all $y \in E_1(\omega) \setminus \{0\}$. Let $\Theta_0(\omega) \in S^{d-1} \cap E_1(\omega)$ be a random variable $\mathcal{F}_0$ measurable and for $t \geq 0$

$$\Theta_t(\omega) = \frac{\varphi(t,\omega)\Theta_0(\omega)}{\|\varphi(t,\omega)\Theta_0(\omega)\|}.$$

We claim that for $\mathbb{P}^J_p$ almost all ω, weak limit points (in the limit $t \to \infty$) of the measures

$$\mu^\omega_t = \frac{1}{t} \int_0^t \delta_{\Theta_s(\omega),\omega_s} ds,\ t \geq 0$$

are invariant measures of the Markov chain (Θ, J). Let μ^ω be such a limit point. Then

$$\lambda_1 = \lim_{t\to\infty} \int \langle A^j\theta, \theta\rangle \mu_t^\omega(d\theta dj) = \Lambda(\mu^\omega).$$

Hence $\lambda_1 = \Lambda(\mu^\omega) \leq \Lambda^+$

The proof of the claim uses classical arguments and goes as follows. Let $C^1(S^{d-1}\times E)$ be the set of maps $g : S^{d-1}\times E \mapsto \mathbb{R}$ that are C^1 in the first variable and let T be the infinitesimal generator of the Markov process (Θ, J). Then T acts on $g \in C^1(S^{d-1}\times E)$ according to a formula similar to (1) with F^i replaced by G^i and $a_{ij}(x)$ replaced by a_{ij}. Because Θ_0 is $\mathcal{F}_0$ measurable,

$$M_t^g = g(\Theta_t(\omega), \omega_t) - g(\Theta_0(\omega), \omega_0) - \int_0^t Tg(\Theta_s(\omega), \omega_s)ds$$

is a $\mathbb{P}_p^J$ martingale with respect to $\{\mathcal{F}_t\}_{t\geq 0}$ where $\mathcal{F}_t = \sigma\{\omega(s) : s \leq t\}$. Its predictable quadratic variation satisfies $\langle M^g\rangle_t = \int_0^t \Gamma(g)(\Theta_s, \omega_s) \leq Ct$ where $\Gamma(g)(\theta, i) = \sum_{ij} a_{ij}(g^j(\theta) - g^i(\theta))^2$. Thus by the strong law of large number for martingales,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t Tg(\Theta_s(\omega), \omega_s)ds = 0 \tag{10}$$

$\mathbb{P}_p^J$ almost surely. Now it is not hard to see that for all $h \geq 0$, $P_h^{\Theta,J}$ preserves $C^1(S^{d-1}\times E)$ (see e.g the proof of Proposition 2.1 in [15]). Thus, we can replace g by $P_h^{\Theta,J}g$ in (10) and since $T(P_h^{\Theta,J}g) = P_h^{\Theta,J}(Tg)$ we get that

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t P_h^{\Theta,J}(Tg)(\Theta_s(\omega), \omega_s)ds = 0$$

for all $h \in \mathbb{Q}^+, \mathbb{P}_p^J$ almost surely. Feller continuity makes $(P_t^{\Theta,J})$ strongly continuous, so that the same is true for all $h \geq 0, \mathbb{P}_p^J$ almost surely. Consequently, $\mathbb{P}_p^J$ almost surely, every weak limit point μ^ω of (μ_t^ω) satisfies $\mu^\omega P_h^{\Theta,J}(Tg) = 0$ for all $h \geq 0$. From the relation $P_t^{\Theta,J}g - g = \int_0^t TP_h^{\Theta,J}gdh$ we then deduce that $\mu^\omega P_t g = \mu^\omega g$ for all $t \geq 0$. Since $C^1(S^{d-1}\times E)$ possesses a countable set dense in $C^0(S^{d-1}\times E)$ it follows that $\mathbb{P}_p^J$ almost surely weak limit points of μ^ω are invariant probabilities of (Θ, J). **QED**

In the multiplicative ergodic theorem, each Lyapunov exponent λ_i comes with an integer $d_i \geq 1$ called its *multiplicity* and such that $\sum_{i=1}^{\tilde{d}} d_i = d$ (see Chapter 3 of [2] for more details). A consequence of Proposition 2.4 is the following inequality which provides, in some cases, a simple way to prove that $\Lambda^+ > 0$, which is often a sufficient condition to ensure positive recurrence of Z on $M \setminus \{0\}\times E$ (see Propostions 2.9 and 2.11 and Theorems 3.2 and 3.3).

Corollary 2.5

$$\sum_{i\in E} p_i \operatorname{Tr}(A^i) = \sum_{i=1}^{\tilde{d}} d_i\lambda_i \leq d\Lambda^+.$$

Proof By Jacobi's formula

$$\frac{\log(\det(\varphi(t,\omega)))}{t} = \frac{\int_0^t \operatorname{Tr}(A^{\omega_s})ds}{t}.$$

By Birkhoff ergodic Theorem, the right hand side of this equality converges, $\mathbb{P}_p^J$ almost surely, as $t \to \infty$, toward $\sum_i p_i \operatorname{Tr}(A^i)$; and a by product of the multiplicative ergodic theorem (see e.g [2], Chapter 3, Corollary 3.3.4) is that the left-hand side converges $\mathbb{P}_p^J$ almost surely, as $t \to \infty$, toward $\sum_{i=1}^{\tilde{d}} d_i \lambda_i$. **QED**

Remark 2.6 If the matrices A^i are *Metzler*, meaning that they have off diagonal nonnegative entries, a result due to Mierczyński ([38], Theorem 1.3) allows to improve the lower bound given in Corollary 2.5 We will use this estimate in section 4, example 4.10.

□

Remark 2.7 Note that in general

$$\Lambda^- \neq \lambda_{\tilde{d}}.$$

Here is a simple example based on [14]. Assume $E = \{1, 2\}$ and $d = 2$ (so that the matrices here are 2×2). Let A^1, A^2 be 2 real matrices having eigenvalues with negative real parts and such that for some $0 < t < 1$, the eigenvalues of $(1-t)A_1 + tA_2$ have opposite signs. It is not hard to construct such a matrix (see e.g [14], Example 1.3). Suppose $a_{12} = \beta t$ and $a_{21} = \beta(1-t)$ with $\beta > 0$, so that $p_1 = (1-t), p_2 = t$. Then, by Corollary 2.5, the Lyapunov exponents, λ_1, λ_2 (counted with their multiplicity) satisfy

$$\lambda_1 + \lambda_2 = (1-t) \operatorname{Tr}(A^1) + t \operatorname{Tr}(A^2) < 0,$$

while, it follows from Theorem 1.6 of [14], that $\Lambda^+ = \Lambda^- > 0$ for β sufficiently large. Hence (for large β)

$$\lambda_2 < 0 < \lambda_1 = \Lambda^- = \Lambda^+.$$

□

2.3 Uniqueness of average growth rate

In this section we discuss general conditions ensuring that

$$\Lambda^- = \Lambda^+ = \lambda_1.$$

A sufficient condition is given by *unique ergodicity* of (Θ, J), meaning that $\mathcal{P}_{inv}^{(\Theta,J)}$ has cardinal one. However, whenever C_M is symmetric (i.e $C_M = -C_M$), for each $\mu \in \mathcal{P}_{inv}^{(\Theta,J)}$ there is another (possibly equal) invariant measure μ^- given as the image measure of μ by the map $x, i \mapsto -x, i$. Indeed, it is easy to see that

$$[\mu P_t^{\Theta,J}]^- = \mu^- P_t^{\Theta,J}$$

for all $\mu \in \mathcal{P}(S^{d-1} \cap C_M \times E)$. This follows from the equivariance property

$$G^i(-x) = -G^i(x)$$

satisfied by the G^i (see equation 5). Clearly $\Lambda(\mu) = \Lambda(\mu^-)$. Thus, when C_M is symmetric, a (weaker than unique ergodicity) sufficient condition is that the quotient space $\mathcal{P}_{erg}^{(\Theta,J)} / \sim$ obtained by identification of μ with μ^- has cardinal one.

Example 2.8 (One dimensional systems) Suppose $d = 1$ and $C_M = \mathbb{R}$. Thus $S^{d-1} \cap C_M = \{\pm 1\}$ and $\mathcal{P}_{erg}^{(\Theta,J)} = \{\mu, \mu^-\}$ where $\mu^i(1) = \mu^{-,i}(-1) = p_i$ and $\mu^i(-1) = \mu^{-,i}(1) = 0$. Hence $\Lambda^- = \Lambda^+ = \lambda_1 = \sum_i p_i a^i$ where $a^i = (F^i)'(0)$.

The two following results complement the previous discussion with practical conditions.

Set $\mathrm{G} = \{G^i\}_{i \in E}, \mathrm{G}_0 = \mathrm{G}, \mathrm{G}_{k+1} = \mathrm{G}_k \cup \{[G^i, V], V \in \mathrm{G}_k\}$ where $[,]$ is the Lie bracket operation. Following [15], we say that the *weak bracket* condition holds at $p \in S^{d-1}$ provided the vector space spanned by the vectors $\{V(p) : V \in \cup_{k \geq 0} \mathrm{G}_k\}$ has full rank (i.e $d-1$).

Proposition 2.9 *Assume there exists $p \in S^{d-1} \cap C_M$ such that*

(i) *The weak bracket condition holds at p;*

(ii) *Either p is G-accessible from $S^{d-1} \cap C_M$ or, C_M is symmetric and $\{-p, p\}$ is G-accessible from $S^{d-1} \cap C_M$.*

Then $\mathcal{P}_{inv}^{(\Theta,J)}$ in the first case, and $\mathcal{P}_{erg}^{(\Theta,J)}/\sim$ in the second, has cardinal one. In particular

$$\Lambda^- = \Lambda^+ = \lambda_1.$$

Proof Existence of an invariant probability follows from compactness and Feller continuity. By Theorem 1 in [5] or Theorem 4.4 in [15] Condition (i), and accessibility of p imply that such a measure is unique (and absolutely continuous with respect to $dx \otimes \sum_i \delta_i$). In case C_M is symmetric and $\{-p, p\}$ accessible, let $S^{d-1} \cap C_M/\sim$ be the projective space obtained by identifying each point x with the antipodal point $-x$ and $\pi : S^{d-1} \cap C_M \mapsto S^{d-1} \cap C_M/\sim$ the quotient map. The PDMP (Θ, J) induces a PDMP $(\pi\Theta, J) = (\pi(\Theta_t), J_t)$ on $S^{d-1} \cap C_M/\sim \times E$ for which $\pi(p)$ is accessible and at which the weak bracket condition holds. The preceding results applies again. **QED**

Example 2.10 (Two dimensional systems) Suppose $d = 2, C_M = \mathbb{R}^2$ and that one of the two following conditions is verified :

(a) At least one matrix, say A^1, has no real eigenvalues; or

(b) at least two matrices, say A^1, A^2 have no (nonzero) common eigenvector.

Then the assumptions, hence the conclusions, of Proposition 2.9 hold.

Indeed, under condition (a), the flow induced by G^1 is periodic on S^1 so that every point $p \in S^1$ satisfies the assumptions of Proposition 2.9. Under condition (b), let $\alpha \leq \beta$ be the eigenvalues of G^1 and $u, v \in S^1$ corresponding eigenvectors. If $\alpha < \beta$ $\{v, -v\}$ is an attractor for the flow induced by G^1 whose basin is $S^1 \setminus \{u, -u\}$. Since $G^2(u) \neq 0$, $\{-v, v\}$ is $\{G^1, G^2\}$ accessible and since $G^2(v) \neq 0$ assumption (i) of Proposition 2.9 is satisfied at point v. If $\alpha = \beta$ every trajectory of the flow induced by G^1 converges either to v or $-v$ and the preceding reasoning still applies.

The next proposition will be useful in Section 4 for analyzing random switching between *cooperative vector fields* and certain epidemiological models. In case the matrices A^i are irreducible, this proposition follows from the Random Perron-Frobenius theorem as proved by

Arnold, Demetrius and Gundlach in [3]. However, to handle the weaker assumption (*iii*), the proof needs to be adapted, but relies on the same ideas. Details are given in Section 7. Recall (see remark 2.6) that a *Metzler* matrix is a matrix with nonnegative off-diagonal entries.

Proposition 2.11 *Assume that*

(i) $C_M = \mathbb{R}^d_+$,

(ii) *For each* $i \in E$, A^i *is Metzler,*

(iii) *There exists* $\alpha \in \mathcal{P}(E)$ *(i.e* $\alpha_i \geq 0, \sum_{i\in E} \alpha_i = 1$*) such that*

$$\overline{A} = \sum_{i\in E} \alpha_i A^i$$

is irreducible.

Then $\mathcal{P}_{inv}^{(\Theta,J)}$ *has cardinal one. In particular*

$$\Lambda^- = \Lambda^+ = \lambda_1.$$

2.4 Average growth rate under frequent switching

The definition of average growth rates (see equations (6) and (7)) involve the invariant measures of (Θ, J) whose explicit computation may prove highly difficult if not impossible. However, when switchings occur frequently, such measures can, by a standard averaging procedure, be estimated by the invariant measures of the mean vector field; i.e the vector field obtained by averaging.

More precisely, we have the following Lemma :

Lemma 2.12 *Assume the switching rates are constant and depend on a small parameter* ε *:* $a^{\varepsilon}_{i,j} = a_{i,j}/\varepsilon$ *where* $(a_{i,j})$ *is an irreducible matrix with invariant probability* p*. Denote by* $(\Theta^{\varepsilon}, J^{\varepsilon})$ *the associated PDMP given by (4), and for any* $\varepsilon > 0$*, let* μ^{ε} *be an element of* $\mathcal{P}_{inv}^{(\Theta^{\varepsilon},J^{\varepsilon})}$*. Then, every limit point of* $(\mu^{\varepsilon})_{\varepsilon>0}$*, in the limit* $\varepsilon \to 0$*, is of the form* $\nu \otimes p$*, where* ν *is an invariant probability measure of the flow induced by* $G^p := \sum_i p_i G^i$*.*

The proof of this lemma follows from standard averaging results. Details are given in Section 7. An immediate corollary is :

Corollary 2.13 *With the hypotheses of Lemma 2.12, assume that the flow induced by* G^p *admits a unique invariant measure* ν *on* $S^{d-1} \cap C_M$*. Denote by* Λ^+_{ε} *and* Λ^-_{ε} *the extremal growth rates of* $(\Theta^{\varepsilon}, J^{\varepsilon})$*. Then*

$$\lim_{\varepsilon\to 0} \Lambda^+_{\varepsilon} = \lim_{\varepsilon\to 0} \Lambda^-_{\varepsilon} = \sum_{i\in E} p_i \int_{S^{d-1}\cap C_M} \langle A^i\theta, \theta\rangle \nu(d\theta).$$

In particular, if $A^p := \sum_i p_i A^i$ *is Metzler and irreducible, then it admits a unique eigenvector* θ^p *on* $S^{d-1} \cap \mathbb{R}^d_+$ *and*

$$\lim_{\varepsilon\to 0} \Lambda^+_{\varepsilon} = \lim_{\varepsilon\to 0} \Lambda^-_{\varepsilon} = \langle A^p\theta^p, \theta^p\rangle = \lambda_{\max}(A^p).$$

3 The non linear system : Main results

3.1 Extinction

The first result is an *extinction* result.

Theorem 3.1 *Assume $\Lambda^+ < 0$. Let $0 < \alpha < -\Lambda^+$. Then there exists a neighborhood $\mathcal{U}$ of 0 and $\eta > 0$ such that for all $x \in \mathcal{U}$ and $i \in E$*

$$\mathbb{P}^Z_{x,i}(\limsup_{t\to\infty} \frac{1}{t}\log(\|X_t\|) \leq -\alpha) \geq \eta.$$

If furthermore 0 is F*-accessible from M, then for all $x \in M$ and $i \in E$*

$$\mathbb{P}^Z_{x,i}(\limsup_{t\to\infty} \frac{1}{t}\log(\|X_t\|) \leq \Lambda^+) = 1.$$

3.2 Persistence

The next results are *persistence* results obtained under the assumption that $\Lambda^- > 0$.

We let

$$\Pi_t = \frac{1}{t}\int_0^t \delta_{Z_s} ds \in \mathcal{P}(M \times E)$$

denote the *empirical occupation measure* of the process Z. For every Borel set $A \subset M \times E$

$$\Pi_t(A) = \frac{1}{t}\int_0^t \mathbf{1}_{\{Z_s \in A\}} ds$$

is then the proportion of the time spent by Z in A up to time t.

We let $M^* = M \setminus \{0\}$.

Theorem 3.2 *Assume $\Lambda^- > 0$. Then the following assertions hold:*

(i) *For all $\varepsilon > 0$ there exists $r > 0$ such that for all $x \in M^*$, $i \in E$, $\mathbb{P}^Z_{x,i}$ almost surely,*

$$\limsup_{t\to\infty} \Pi_t(B(0,r) \times E) \leq \varepsilon.$$

In particular, for all $x \in M^*$, $\mathbb{P}^Z_{x,i}$ almost surely, every limit point (for the weak topology) of (Π_t) belongs to $\mathcal{P}^Z_{inv} \cap \mathcal{P}(M^* \times E)$.*

(ii) *There exist positive constants θ, K such that for all $\mu \in \mathcal{P}^Z_{inv} \cap \mathcal{P}(M^* \times E)$*

$$\sum_{i\in E} \int \|x\|^{-\theta} \mu^i(dx) \leq K.$$

(iii) *Let $\varepsilon > 0$ and τ^ε be the stopping time defined by*

$$\tau^\varepsilon = \inf\{t \geq 0 : \|X_t\| \geq \varepsilon\}.$$

There exist $\varepsilon > 0$, $b > 1$ and $c > 0$ such that for all $x \in M^*$ and $i \in E$,

$$\mathbb{E}^Z_{x,i}(b^{\tau^\varepsilon}) \leq c(1 + \|x\|^{-\theta}).$$

Set $\mathrm{F}_0 = \mathrm{F} = \{F^i\}_{i\in E}$ and $\mathrm{F}_{k+1} = \mathrm{F}_k \cup \{[F^i, V], V \in \mathrm{F}_k\}$ where $[,]$ is the Lie bracket operation. We say (compare to Section 2.3) that the *weak bracket* condition holds at $p \in M$ provided the vector space spanned by the vectors $\{V(p) : V \in \cup_{k\geq 0}\mathrm{F}_k\}$ has full rank. We let **Leb** denote the Lebesgue measure on $\mathbb{R}^d$.

Theorem 3.3 *In addition to the assumption $\Lambda^- > 0$, assume that there exists a point $p \in M^*$ F-accessible from M^* at which the weak bracket condition holds. Then*

(i)

$$\mathcal{P}^Z_{inv} \cap \mathcal{P}(M^* \times E) = \{\Pi\};$$

(ii) Π *is absolutely continuous with respect to* $\mathbf{Leb} \otimes (\sum_{i\in E} \delta_i)$

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

$$\lim_{t\to\infty} \Pi_t = \Pi$$

$\mathbb{P}^Z_{x,i}$ *almost surely.*

In order to get a convergence in distribution, the weak bracket condition needs to be strengthened. Set $\mathcal{F}_0 = \{F^i - F^j : i,j = 1,\dots m\}$ and $\mathcal{F}_{k+1} = \mathcal{F}_k \cup \{[F^i, V] : V \in \mathcal{F}_k\}$. We say that the *strong bracket* condition holds at $p \in M$ provided the vector space spanned by the vectors $\{V(p) : V \in \cup_{k\geq 0}\mathcal{F}_k\}$ has full rank.

Given $\mu, \nu \in \mathcal{P}(M \times E)$, the *total variation distance* between μ and ν is defined as

$$\|\mu - \nu\|_{TV} = \sup |\mu(A) - \nu(A)|$$

where the supremum is taken over all Borel sets $A \subset M \times E$.

Theorem 3.4 *Under the assumptions of the preceding theorem, if the weak Bracket condition is strengthened to the strong bracket condition, then there exist $\kappa, \theta > 0$ such that for all $x \in M^*$ and $i \in E$,*

$$\|\mathbb{P}^Z_{x,i}(Z_t \in \cdot) - \Pi\|_{TV} = \|\delta_{x,i}P^Z_t - \Pi\|_{TV} \leq const.(1 + \|x\|^{-\theta})e^{-\kappa t}.$$

4 Epidemic Models in Fluctuating Environment

We discuss here some implications of our results to certain epidemics models evolving in a randomly fluctuating environment.

Forty years ago, Lajmanovich and Yorke in a influential paper [33], proposed and analyzed a deterministic SIS (susceptible-infectious-susceptible) model of infection, describing the evolution of a disease that does not confer immunity, in a population structured in d groups. The model is given by a differential equation on $[0,1]^d$ (the unit cube of $\mathbb{R}^d$) having the form

$$\frac{dx_i}{dt} = (1 - x_i)(\sum_{j=1}^d C_{ij}x_j) - D_i x_i\,,\ i = 1,\dots d, \tag{11}$$

where $C = (C_{ij})$ is an irreducible matrix with nonnegative entries and $D_i > 0$. Here $0 \leq x_i \leq 1$ represents the proportion of infected individuals in group i; D_i is the intrinsic cure rate in group

i and $C_{ij} \geq 0$ is the rate at which group i transmits the infection to group j. Irreducibility of C implies that each group indirectly affects the other groups. By a classical mean field approximation procedure, (11) can be derived from a finite population model, in the limit of an infinite population (see Benaïm and Hirsch [11]).

Here and throughout, for any matrix A we let $\lambda(A)$ denote the largest real part of the eigenvalues of A. A matrix A is called *Hurwitz* provided $\lambda(A) < 0$. Lajmanovich and Yorke [33] prove the following result:

Theorem 4.1 (Lajmanovich and Yorke, [33]) *Let* $A = C - \mathsf{diag}(D)$.

If $\lambda(A) \leq 0$, 0 *is globally asymptotically stable for the semiflow induced by (11) on* $[0,1]^d$.

If $\lambda(A) > 0$ *there exists another equilibrium* $x^* \in]0,1[^d$ *whose basin of attraction is* $[0,1]^d \setminus \{0\}$.

In this epidemiological framework, 0 is called the *disease free equilibrium*, and the point x^*, when it exists, the *endemic equilibrium*. It turns out that such a dichotomic behavior is very robust to the perturbations of the model and can be obtained under a very general set of assumptions, using Hirsch's theory of *cooperative differential equations*.

We let $\mathbb{R}^d_{++}$ denote the interior of the non negative orthant $\mathbb{R}^d_+$. For $x, y \in \mathbb{R}^d$ we write $x \leq y$ (or $y \geq x$) if $y - x \in \mathbb{R}^d_+$; $x < y$ if $x \leq y$ and $x \neq y$; and $x << y$ if $y - x \in \mathbb{R}^d_{++}$.

Following [11] (especially Section 3), we call a map $F : [0,1]^d \mapsto \mathbb{R}^d$ an *epidemic vector field* if it is C^1 and[1] satisfies the following set of conditions:

E1 $F(0) = 0$;

E2 $x_i = 1 \Rightarrow F_i(x) < 0$;

E3 F is *cooperative* i.e the Jacobian matrix $DF(x)$ is Metzler for all $x \in [0,1]^d$;

E4 F is *irreducible* on $[0,1)^d$ i.e $DF(x)$ is irreducible for all $x \in [0,1)^d$;

E5 F is strongly *sub-homogeneous* on $(0,1)^d$ i.e $F(\lambda x) << \lambda F(x)$ for all $\lambda > 1$ and $x \in (0,1)^d$.

It is easy to verify that the *Lajmanovich and Yorke vector field* (given by the right hand side of (11)) satisfies these conditions.

Let $\Psi = \{\Psi_t\}$ denote the local flow induced by F. Condition $E3$ has the important consequence that for all $t \geq 0$ Ψ_t is *monotone* for the partial ordering $\leq$. That is $\Psi_t(x) \leq \Psi_t(y)$ if $x \leq y$. In particular, by $E1$, $\Psi_t(x) \geq 0$ for all $x \geq 0$. Combined with $E2$ this shows that $[0,1]^d$ is positively invariant under Ψ.

The following result shows that trajectories of Ψ behave exactly like the trajectories of the Lajmanovich and Yorke system. The first assertion was stated in ([11], Theorem 3.2) but its proof is a consequence of more general results due to Hirsch (in particular Theorems 3.1 and 5.5 in [30]).

Theorem 4.2 *Let F be an epidemic vector field and $\Psi = \{\Psi_t\}_{t\geq 0}$ the induced semiflow on* $[0,1]^d$. *Then*

(i) *(Hirsch, [30]) Either* 0 *is globally asymptotically stable for* Ψ; *or there exists another equilibrium* $x^* \in]0,1[^d$ *whose basin of attraction is* $[0,1]^d \setminus \{0\}$.

[1] by this we mean that F can be extended to a C^1 vector field on $\mathbb{R}^d$.

(ii) *Let $A = DF(0)$. Then 0 is globally asymptotically stable if and only if $\lambda(A) \leq 0$.*

Proof As already mentioned, (i) follows from [30], Theorems 3.1 and 5.5. We detail the proof of (ii). If $\lambda(A) < 0$, then 0 is linearly stable hence globally stable by (i). If $\lambda(A) > 0$, there exists, by irreducibility and Perron Frobenius theorem, $x_0 >> 0$ such that $Ax_0 = \lambda(A)x_0 >> 0$. Hence $F(\varepsilon x_0) >> 0$ for ε small enough, because $\frac{F(\varepsilon x_0)}{\varepsilon} \to Ax_0$ as $\varepsilon \to 0$. Consequently $\{x : x \geq \varepsilon x_0\}$ is positively invariant and 0 cannot be asymptotically stable.

It remains to show that 0 is asymptotically stable when $\lambda(A) = 0$. Suppose the contrary. By (i) there exists another equilibrium $x^* >> 0$. Set $y^* = x^*/2$. By strong subhomogeneity, $0 = F(x^*) << 2F(y^*)$. Let $F_\varepsilon(x) = F(x) - \varepsilon x$. For all $\varepsilon > 0$, F_ε is an epidemic vector field and 0 is linearly stable for F_ε (because $\lambda(DF_\varepsilon(0)) = -\varepsilon$). On the other hand, for ε small enough, $0 << F_\varepsilon(y^*)$ so that the set $\{y : y \geq y^*\}$ is positively invariant by F_ε. A contradiction. **QED**

4.1 Fluctuating environment

We consider a PDMP $Z = (X, I)$ as defined in Section 1, under the assumptions that:

E'1 $M = [0, 1]^d$;

E'2 For all $i \in E$, $A^i = DF^i(0)$ is Metzler;

E'3 There exists $\alpha \in \mathcal{P}(E)$ such that the convex combination $\overline{A} = \sum_{i \in E} \alpha_i A^i$ is irreducible.

Observe that these conditions are automatically satisfied if $\mathrm{F} = \{F^i\}_{i \in E}$ consists of epidemic vector fields but are clearly much weaker.

Relying on Proposition 2.11, we let $\lambda_1 = \Lambda^+ = \Lambda^-$ denote the top Lyapunov exponent of the linearized system.

Theorem 4.3 *Assume $\lambda_1 < 0$ and that one of the following two conditions holds:*

(a) *The jump rates are constant (i.e $a_{ij}(x) = a_{ij}$) and the F^i are epidemic; or*

(b) *There exists $\beta \in \mathcal{P}(E)$ such that $\overline{F} = \sum_i \beta_i F^i$ is epidemic and*

$$\lambda(\sum_i \beta_i A^i) \leq 0.$$

Then for all $x \in M^*$ and $i \in E$,

$$\mathbb{P}^Z_{x,i}(\limsup \frac{\log(\|X_t\|)}{t} \leq \lambda_1) = 1.$$

Proof We first prove the result under condition (a). Recall (see Section 2.2) that Ω stands for $D(\mathbb{R}^+, E)$. For each $\omega \in \Omega$ and $x \in [0, 1]^d$ let

$$t \mapsto \Psi(t, \omega)(x)$$

be the solution to the non autonomous differential equation

$$\dot{y} = F^{\omega_t}(y),$$

with initial condition $y(0) = x$. By conditions $E3$ and $E5$ each flow Ψ^i is monotone and subhomogenous (see e.g [30], Theorem 3.1). The composition of monotone subhomogeneous mappings being monotone and subhomogeneous, $\Psi(t, \omega)$ is monotone and subhomogeneous for all $t \geq 0$ and $\omega \in \Omega$. Thus, for all $\varepsilon > 0$ and $\|x\| > \varepsilon$

$$\Psi(t, \omega)(x) \leq \frac{\|x\|}{\varepsilon} \Psi(t, \omega)(\frac{\varepsilon}{\|x\|} x). \tag{12}$$

Under the assumption that the jump rates are constant, $\mathbb{P}^Z_{x,i}$ is the image measure of $\mathbb{P}^J_i$ by the map

$$\omega \mapsto (\omega, (\Psi(t, \omega)(x))_{t \geq 0}).$$

Therefore, by Theorem 3.1, there exists $\eta, \varepsilon > 0$ such that for all $x \in B(0, \varepsilon)$

$$\mathbb{P}^Z_{x,i}(\limsup_{t \to \infty} \frac{\log(\|X_t\|)}{t} \leq \lambda_1) = \mathbb{P}^J_i(\limsup_{t \to \infty} \frac{\log(\|\Psi(t, \omega)(x)\|)}{t} \leq \lambda_1) \geq \eta. \tag{13}$$

Combined with (12), this proves that (13) holds true not only for $x \in B(0, \varepsilon)$ but for all $x \in [0, 1]^d$. A standard application of the Markov property then implies the result.

Under condition (b), it follows from Theorem 4.2, that 0 is F-accessible from M, and the result follows from Theorem 3.1. **QED**

Remark 4.4 The assumption made in case (a) that the F^i are epidemic can be weakened. The proof shows that irreducibility of F^i is unnecessary and that strong subhomogeneity can be weakened to subhomogeneity.

□

Remark 4.5 Case (a) (and its proof) can be related with the results obtained by Chueshov in [19], for SIS models with random coefficients (see [19, Section 5.7.2]) and, more generally, for monotone subhomogeneous random dynamical systems. Note, however, that in comparison with Chueshov's approach, in case (b), there is no assumption that the F^is are monotone nor subhomogeneous.

□

Example 4.6 (Fluctuations may promote cure) We give here a simple example consisting of two Lajmanovich-Yorke vector fields modeling the evolution of an endemic disease (each vector field possesses an endemic equilibrium) but such that a random switching between the dynamics leads to the extinction of the disease.

Suppose $d = 2, E = \{0, 1\}$. Let F^0, F^1 be the Lajmanovich-Yorke vector fields respectively given by

$$C^0 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \ D^0 = \begin{pmatrix} 6 \\ 1 \end{pmatrix},$$

and

$$C^1 = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}, \ D^1 = \begin{pmatrix} 1 \\ 7 \end{pmatrix}.$$

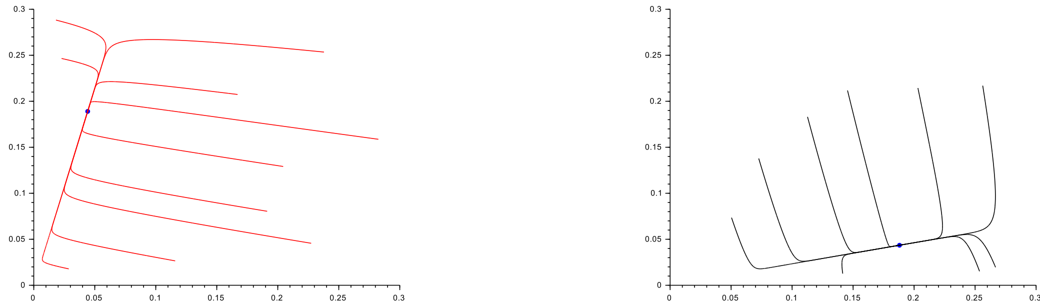

Figure 1: Example 4.6, phase portrait of F^0 and F^1

One can easily check that

$$\lambda(A^0) = \lambda(A^1) = 4 + \sqrt{20} > 0,$$

so that for each F^i, there is an endemic equilibrium and the disease free equilibrium is a repellor. On the other hand,

$$\lambda(\frac{A^0 + A^1}{2}) = -2 < 0,$$

so that the disease free equilibrium is a global attractor of the average vector field $\overline{F} = \frac{1}{2}(F^1 + F^2)$. Consider now the PDMP given by constant switching rates

$$a_{0,1} = a_{1,0} = \beta, a_{0,0} = a_{1,1} = 0.$$

By Corollary 2.13, this implies that $\lambda_1 < 0$ provided β is sufficiently large. Thus the conclusion of Theorem 4.3 holds.

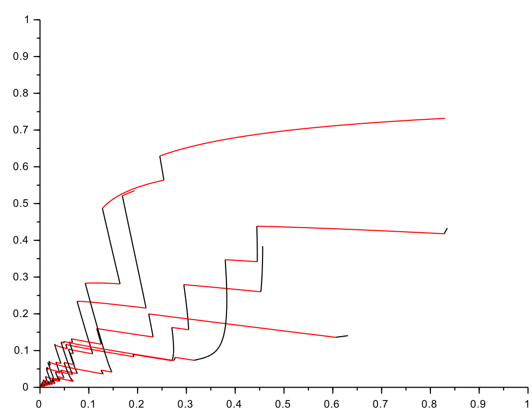

Figure 2: Example 4.6, some trajectories of (X_t) for $\beta = 20$

Example 4.7 (Fluctuations may promote infection) We give here another simple example consisting of two Lajmanovich-Yorke vector fields for which the disease dies out, but such that a random switching between the dynamics leads to the persistence of the disease.

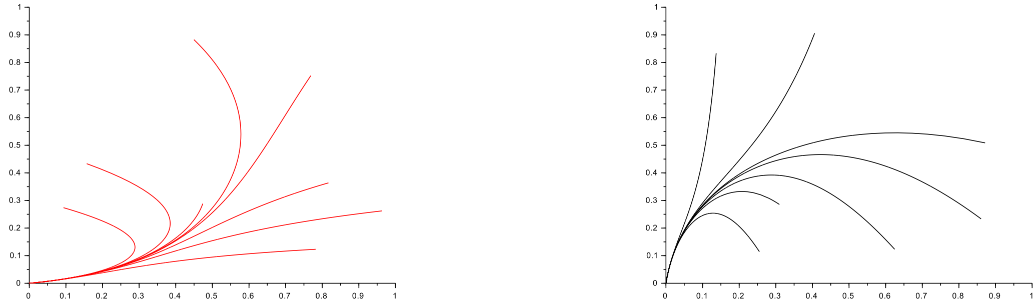

Figure 3: Example 4.7, Phase portrait of F^0 and F^1

With the notation of Example 4.6, assume now that

$$C^0 = \begin{pmatrix} 1 & 4 \\ \frac{1}{16} & 1 \end{pmatrix},\ D^0 = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

and

$$C^1 = \begin{pmatrix} 2 & \frac{1}{16} \\ 4 & 2 \end{pmatrix},\ D^1 = \begin{pmatrix} 3 \\ 3 \end{pmatrix}.$$

Straightforward computation shows that

$$\lambda(A^0) = \lambda(A^1) = -1/2 < 0,$$

$$\lambda(\frac{A^0 + A^1}{2}) = 33/32 > 0,$$

and that the endemic equilibrium of $\overline{F}$ is the point $x^\star = (33/113, 33/113)$. Then $x^\star$ is F -accessible and one can easily check that the strong bracket condition holds at $x^\star$. Thus, for β sufficiently large, this implies by Corollary 2.13 and Theorem 3.4 the exponential convergence in total variation of the distribution of Z_t (whenever $X_0 \neq 0$) towards a unique distribution Π absolutely continuous with respect to $\mathsf{Leb} \otimes \sum_{i \in E} \delta_i$ and satisfying the tail condition given by Theorem 3.2 (ii). Furthermore, it follows from ([15], Proposition 3.1) that the topological support of Π writes $\Gamma \times E$ where Γ is a compact connected set containing both 0 and $x^\star$, and whose interior is dense in Γ.

Remark 4.8 In the preceding example, the matrices A^i are Metzler and Hurwitz but $\lambda_1 > 0$ because the convex hull of the $\{A^i\}$ contains a non Hurwitz matrix. This leads to the natural question of finding examples for which:

$\lambda_1 > 0$ and every matrix in the convex hull of the $\{A^i\}$ is Hurwitz.

For arbitrary (i.e non Metzler) matrices, such and example has been given in dimension 2 in [34] and more recently in [32].

Now, if we restrain ourselves to Metzler matrices, a result from Gurvits, Shorten and Mason ([26, Theorem 3.2]) proves that, in dimension 2, when every matrix in the convex hull is Hurwitz, then 0 is globally asymptotically stable for any deterministic switching between the linear systems. In particular, this implies that λ_1 cannot be positive.

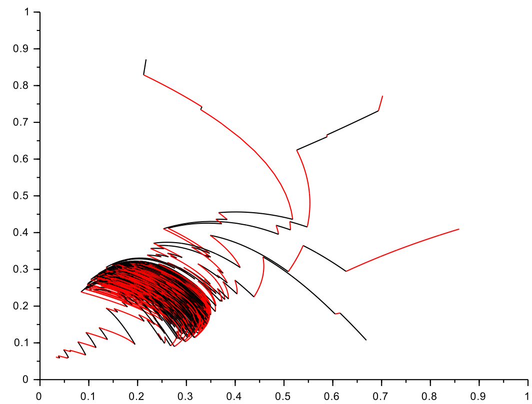

Figure 4: Example 4.7, some trajectories of (X_t) for $\beta = 20$

However, they show that it is possible in some higher dimension to construct an example where all the matrices in the convex hull are Hurwitz, and for which there exists a periodic switching such that the linear system explodes. Later, an explicit example in dimension 3 was given by Fainshil, Margaliot and Chiganski [23]. Precisely, consider the matrices

$$A^0 = \begin{pmatrix} -1 & 0 & 0 \\ 10 & -1 & 0 \\ 0 & 0 & -10 \end{pmatrix}, \; A^1 = \begin{pmatrix} -10 & 0 & 10 \\ 0 & -10 & 0 \\ 0 & 10 & -1 \end{pmatrix}.$$

It is shown in [23] that every convex combination of A^0 and A^1 is Hurwitz, and yet a switch of period 1 between A^0 and A^1 yields an explosion. Some simulations made on Scilab (see Figure 5) let us think that this result is still true for a random switching, with rates

$$a_{0,1} = a_{1,0} = \beta, a_{0,0} = a_{1,1} = 0.$$

Here β has to be chosen neither too small nor too big. Using the formula

$$\lim_{t\to\infty} \mathbb{E}(\frac{1}{t}\int_0^t \langle A^{J_s}\Theta_s, \Theta_s\rangle ds) = \lambda_1(\beta),$$

and Monte-Carlo simulations we can estimate numerically $\lambda_1(\beta)$. The results are plotted in Figure 6 and show (although we didn't prove it) that $\lambda_1 > 0$ for $3 \le \beta \le 30$, providing a positive answer to the question raised at the beginning of the remark.

□

Example 4.9 (Fluctuations may promote infection, continued) Remark 4.8 can be used to produce two Lajmanovich-Yorke vector fields F^0, F^1 on $[0,1]^3$ such that

(i) For all $0 \le t \le 1$, the disease free equilibrium is a global attractor of the vector field $F^t = (1-t)F^0 + tF^1$;

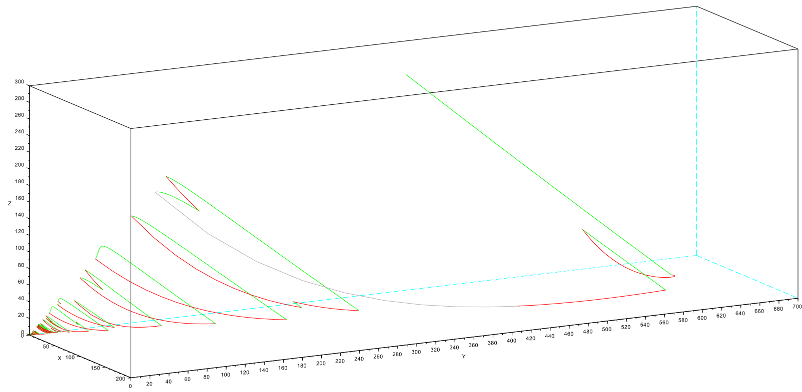

Figure 5: Simulation of Y_t for $\beta = 10$.

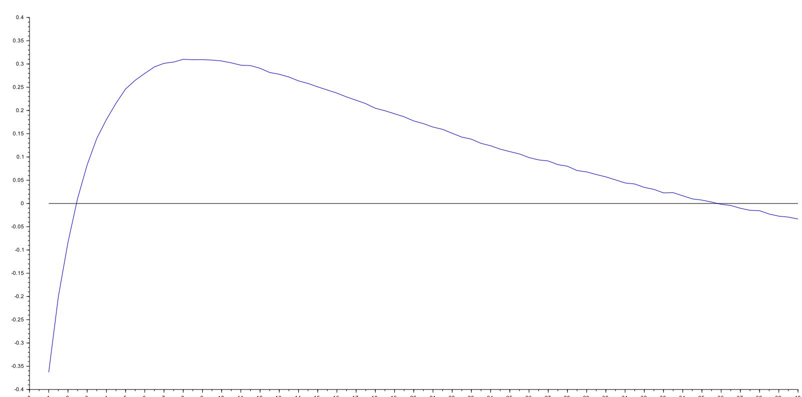

Figure 6: Approximation of $\lambda_1(\beta)$ by Monte-Carlo method

(ii) A random switching between the dynamics leads to the persistence of the disease.

Observe that F^t is the Lajmanovich-Yorke vector field with infection matrix $C^t = (1-t)C^0 + tC^1$ and cure rate vector $D^t = (1-t)D^0 + tD^1$

To do so, one just has to choose C^0, C^1, D^0, D^1 in such way that $A^i = C^i - D^i$. For the simulation given here, we have chosen

$$D^0 = \begin{pmatrix} 11 \\ 11 \\ 20 \end{pmatrix},$$

and

$$D^1 = \begin{pmatrix} 20 \\ 20 \\ 11 \end{pmatrix}.$$

When (see Figure 6) β is such that $\lambda_1 > 0$, then by Theorem 4.11 below, Z admits a unique invariant measure Π on $M^* \times E$. Moreover by Theorem 3.2, there exists $\theta > 0$ such that

$$\sum_{i \in E} \int \|x\|^{-\theta} \Pi^i(dx) < \infty.$$

Figure 7 and 8 illustrate this persistence of the infection.

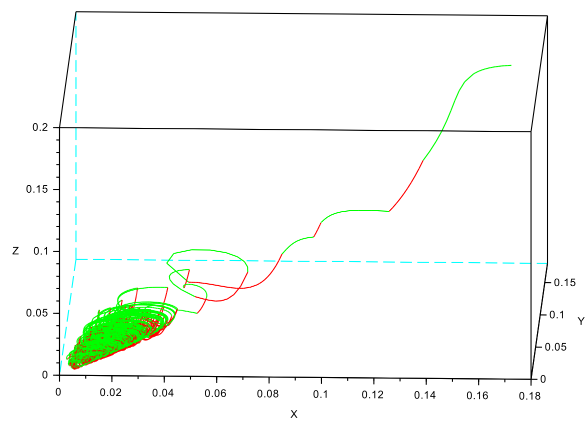

Figure 7: Example 4.9 : Simulation of X_t for $\beta = 10$.

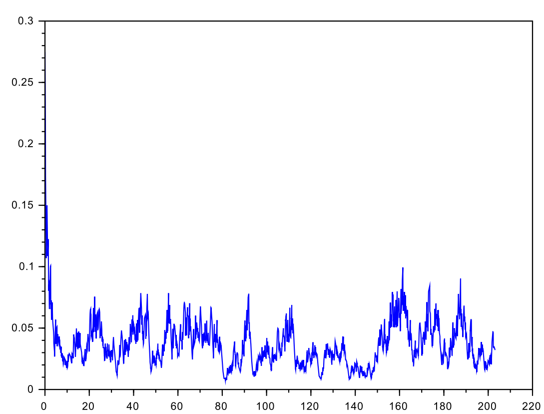

Figure 8: Example 4.9 : Simulation of $\|X_t\|$ for $\beta = 10$.

4.2 Exponential convergence without bracket condition

Throughout with section, we assume that the vector fields F^i are epidemic and that the jump rates are constant. Recall (see proof of Theorem 4.3) that this implies that for all $\omega \in \Omega$ and $t > 0$, $\Psi(t, \omega)$ is monotone and strongly subhomegeneous. A very useful consequence of this fact is the strict nonexpansivity of $\Psi(t, \omega)$ on $\mathbb{R}^d_{++}$ with respect to the Birkhoff part metric p, the definition of which is recalled below. Now if we assume that $\lambda_1 > 0$, we have a Lyapunov function and nonexpansivity, so we might except uniqueness of the invariant measure on $[0,1]^d \setminus \{0\} \times E$ and convergence in law of (Z_t) towards it. Here we prove that this is indeed the case, and even that we have an exponential rate of convergence towards this invariant measure for a certain Wasserstein distance, thanks to a weak form of Harris' theorem given by Hairer, Mattingly and Scheutzow [27]. But before to do so, we explain briefly why we cannot expect to have convergence in total variation without additional assumptions with the following simple example :

Example 4.10 Suppose $d = 2, E = \{0, 1\}$. Let F^0, F^1 be the *Lajmanovich-Yorke vector fields* respectively given by

$$C^0 = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \; D^0 = \begin{pmatrix} 2 \\ 3 \end{pmatrix},$$

and

$$C^1 = \begin{pmatrix} 6 & 2 \\ 7 & 3 \end{pmatrix}, \; D^1 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

One can easily check that the point $x^* = (1/2, 1/2)$ is a common equilibrium of F^1 and F^2. In particular, $\Pi = \delta_{x^*} \otimes (\delta_0 + \delta_1)/2$ is an invariant probability of Z. Moreover, for all $x \neq x^*$, $i \in E$ and $t \geq 0$, one has $\mathbb{P}^Z_{x,i}(Z_t \in \{x^*\} \times E) = 0$ so $\|\delta_{x,i} P^Z_t - \Pi\|_{TV} = 1$ for all $t \geq 0$. Now let us quickly show that X_t converges almost surely exponentially fast to x^*, for all switching rates. Let $\lambda_1(0) = \lambda_1$ (respectively $\lambda_1(x^*)$) denote the top Lyapunov exponent of the linearized system at the origin (respectively at x^*). By Proposition 2.11 this exponent coincides with the unique average growth rate of the corresponding linearized system. We claim that $\lambda_1(0) > 0$ and $\lambda_1(x^*) < 0$. The first inequality follows from the Kolotilina-type lower estimate for the top Lyapunov exponent mentioned in Remark 2.6 due to Mierczyński ([38, Theorem 1.3]). In our setting, this estimate ensures that

$$\lambda_1(0) \geq \frac{1}{2} \sum_i p_i Tr(A^i) + \sum_i p_i \sqrt{A^i_{12} + A^i_{21}},$$

which is positive because $Tr(A^0) = Tr(A^1) = 0$ and the other terms are positive. Let $B^i = DF^i(x^*)$. Then the second estimate follows from Lemma 2.3 because one can easily check that $\lambda_{\max}(B^1 + (B^1)^T) \leq \lambda_{\max}(B^0 + (B^0)^T) < 0$. So applying Theorem 3.1, we have a neighborhood $\mathcal{U}$ of x^* and $\eta > 0$ such that for all $x \in \mathcal{U}$ and $i \in E$

$$\mathbb{P}^Z_{x,i}(\limsup_{t \to \infty} \frac{1}{t} \log(\|X_t - x^*\|) \leq \frac{\lambda_1(x^*)}{2}) \geq \eta. \tag{14}$$

On the other hand, because $\lambda_1(0) > 0$, there exists by Theorem 3.2 $\varepsilon > 0$ such that for all $x \neq 0$,

$$\mathbb{P}^Z_{x,i}(\tau < \infty) = 1, \tag{15}$$

where $\tau = \inf\{t \geq 0 \ : \ \|X_t| \ \geq \varepsilon\}$. Finally, because x^* is a linear stable equilibrium for F^0 with basin of attraction contains M^*, one can show that there exists a constant $c > 0$ such that for all $x \in M$ with $\|x\| \geq \varepsilon$,

$$\mathbb{P}^Z_{x,i}(Z_t \in \mathcal{U} \times E) \geq c. \tag{16}$$

Combining (14), (15), (16) and the Markov property implies that

$$\mathbb{P}^Z_{x,i}(\limsup_{t\to\infty} \frac{1}{t} \log(\|X_t - x^*\|) \leq \lambda_1(x^*)) = 1,$$

for all $(x, i) \in M^* \times E$ (see [16, Theorem 3.1] for details on a very similar proof).

Before stating our theorem, recall the definition of the Wasserstein distance. Let $\mathcal{Y}$ be a polish space, and d be a distance-like function on $\mathcal{Y}$. That is d satisfies the axioms of a distance, except for the triangle inequality. Then the *Wasserstein distance* associated to d is defined for every $\mu, \nu \in \mathcal{P}(\mathcal{Y})$ by

$$\mathcal{W}_d(\mu, \nu) = \inf_{\pi \in C(\mu,\nu)} \int_{\mathcal{X}^2} d(x, y) \mathrm{d}\pi(x, y),$$

where $C(\mu, \nu)$ is the set of all the coupling of μ and ν. When d is a distance, so is $\mathcal{W}_d$, and in every case, $\mathcal{W}_d(\mu, \nu) = 0$ if and only if $\mu = \nu$.

Set $\mathcal{Y} = [0, 1]^d \setminus \{0\} \times E$.

Theorem 4.11 *Assume the F^i are epidemic vector fields, (a_{ij}) are constant and $\lambda_1 > 0$. Then there exists a distance-like function $\tilde{d}$, $t_0 \geq 0$ and $r > 0$, such that,*

(i) *for all $t \geq t_0$, for all $\mu, \nu \in \mathcal{P}(\mathcal{X})$,*

$$\mathcal{W}_{\tilde{d}}(\mu P^Z_t, \nu P^Z_t) \leq e^{-rt} \mathcal{W}_{\tilde{d}}(\mu, \nu).$$

(ii) *(P^Z_t) has a unique invariant measure Π on $\mathcal{X}$, and for all $\mu \in \mathcal{P}(\mathcal{X})$,*

$$\mathcal{W}_{\tilde{d}}(\mu P^Z_t, \Pi) \leq e^{-rt} \mathcal{W}_{\tilde{d}}(\mu, \Pi).$$

5 Proofs of Theorems 3.1–3.4 : A *stochastic persistence* approach

As indicated in the introduction, the proofs will be deduced from the qualitative properties of PDMPs combined with general results on *stochastic persistence* proved in [9] along the lines of the seminal results obtained by Schreiber, Hofbauer and their co-authors.

5.1 An abstract stochastic persistence result

The results in [9] concern certain Markov processes on a (possibly) non compact metric space satisfying a weak version of the Feller property. Here for simplicity we shall state a simpler version of these results tailored for Feller processes on a compact space.

Let $\mathcal{X}$ be a **compact** metric space and $\tilde{Z}$ a cad-lag Markov process on $\mathcal{X}$. To shorten notation we write $\mathbb{P}_x, \mathbb{P}_\mu, (P_t)_{t\geq 0}, \mathcal{P}_{inv}, \mathcal{P}_{erg}$ in place of $\mathbb{P}_x^{\tilde{Z}}, \mathbb{P}_\mu^{\tilde{Z}}, (P_t^{\tilde{Z}})_{t\geq 0}, \mathcal{P}_{inv}^{\tilde{Z}}, \mathcal{P}_{erg}^{\tilde{Z}}$. We let

$$\Pi_t = \frac{1}{t}\int_0^t \delta_{\tilde{Z}_s} ds$$

denote the empirical occupation measure of $\tilde{Z}$. We let $C(\mathcal{X})$ denotes the space of real valued continuous functions on $\mathcal{X}$ equipped with the uniform norm $\|f\| = \sup_{x\in\mathcal{X}} |f(x)|$.

We assume that $(P_t)_{t\geq 0}$ is *Feller*. That is

(a) For all $t \geq 0$ P_t maps $C(\mathcal{X})$ into itself,

(b) For all $f \in C(\mathcal{X})$ $\lim_{t\to 0} \|P_t f - f\| = 0$.

We let $\mathcal{L}$ denote the *infinitesimal generator* of (P_t) and $\mathcal{D}$ its domain. Recall that $\mathcal{D}$ is defined as the set of $f \in C(\mathcal{X})$ such that $\frac{1}{t}(P_t f - f)$ converges in $C(\mathcal{X})$, and, for such an f, $\mathcal{L}f$ denotes the limit. We let $\mathcal{D}^2 \subset \mathcal{D}$ denote the set of $f \in \mathcal{D}$ such that $f^2 \in \mathcal{D}$. For $f \in \mathcal{D}^2$ the *Carré du champ* of f is defined as

$$\Gamma(f) = \mathcal{L}f^2 - 2f\mathcal{L}f. \tag{17}$$

We assume that

Hypothesis 5.1 *there exists a non empty compact set $\mathcal{X}_0 \subset \mathcal{X}$ called the* extinction set *which is invariant under $(P_t)_{t\geq 0}$. That is*

$$P_t \mathbb{1}_{\mathcal{X}_0} = \mathbb{1}_{\mathcal{X}_0}$$

where $\mathbb{1}_{\mathcal{X}_0}$ stands for the indicator function of $\mathcal{X}_0$.

We set

$$\mathcal{X}_+ = \mathcal{X} \setminus \mathcal{X}_0,$$

$\mathcal{P}_{inv}(\mathcal{X}_+) = \mathcal{P}_{inv} \cap \mathcal{P}(\mathcal{X}_+), \mathcal{P}_{inv}(\mathcal{X}_0) = \mathcal{P}_{inv} \cap \mathcal{P}(\mathcal{X}_0)$ etc.

Extinction of $\tilde{Z}$ amounts to say that trajectories of $(\tilde{Z}_t)$ converge almost surely to $\mathcal{X}_0$. Let $\mathcal{X}_0^\varepsilon$ be the ε-neighborhood of $\mathcal{X}_0$. Using a terminology borrowed to Schreiber [41] and Chesson [18], we say that $\tilde{Z}$ is *stochastically persistent* (or *almost surely persistent*), respectively *persistent in probability*, provided

$$\lim_{\varepsilon\to 0} \limsup_{t\to\infty} \Pi_t(\mathcal{X}_0^\varepsilon) = 0$$

$\mathbb{P}_x$ almost surely for all $x \in \mathcal{X}_+$. Respectively

$$\lim_{\varepsilon\to 0} \limsup_{t\to\infty} \mathbb{P}_x(Z_t \in \mathcal{X}_0^\varepsilon) = 0$$

for all $x \in \mathcal{X}_+$.

General criteria ensuring extinction or persistence are given by the existence of a suitable *average Lyapounov function* V as defined now.

In addition to hypothesis 5.1 we assume that

Hypothesis 5.2 *There exist continuous maps $V : \mathcal{X}_+ \mapsto \mathbb{R}^+$ and $H : \mathcal{X} \mapsto \mathbb{R}$ enjoying the following properties :*

(a) *For all compact $K \subset \mathcal{X}_+$ there exists $V_K \in \mathcal{D}^2$ with $V|_K = V_K|_K$ and $(\mathcal{L}V_K)|_K = H|_K$;*

(b) $\sup_{\{K: K\subset\mathcal{X}_+, K \text{ compact }\}} \|\Gamma(V_K)|_K\| < \infty$;

(c) $\lim_{x\to\mathcal{X}_0} V(x) = \infty$;

(d) *Jumps of $V(\tilde{Z}_t)$ are bounded : $\exists\Delta > 0$ such that $|V(\tilde{Z}_t) - V(\tilde{Z}_{t-})| \leq \Delta$;*

(e) *For all $t \geq 0$ $P_t H$ and $\int_0^t P_s H ds$ lie in $\mathcal{D}^2$.*

Let $\mathcal{P}_{erg}(\mathcal{X}_0) = \mathcal{P}_{erg} \cap \mathcal{P}(\mathcal{X}_0)$. Define the H-exponents of the processes as

$$H_* = \inf_{\mu\in\mathcal{P}_{erg}(\mathcal{X}_0)} \mu H \text{ and } H^* = \sup_{\mu\in\mathcal{P}_{erg}(\mathcal{X}_0)} \mu H.$$

We call the process *H-persistent* if $H^* < 0$ and *H-nonpersistent* if $H_* > 0$.

By the Ergodic decomposition theorem, note that $H^* < 0$ (respectively $H_* > 0$) if and only if $\mu H < 0$ (respectively > 0) for all $\mu \in \mathcal{P}_{erg}(\mathcal{X}_0)$.

We say that $A \subset \mathcal{X}$ is *accessible* from $B \subset \mathcal{X}$ if for every neighborhood U of A and $x \in B$ there exists $t \geq 0$ such that $P_t \mathbb{1}_U(x) > 0$.

Theorem 5.3 *Suppose that the process is H-nonpersistent. Then*

(i) *For all $0 < \alpha < H_*$, there exists a neighborhood U of $\mathcal{X}_0$ and $\eta > 0$ such that*

$$\mathbb{P}_x(\liminf_{t\to\infty} \frac{V(\tilde{Z}_t)}{t} \geq \alpha) \geq \eta$$

for all $x \in U$;

(ii) *If furthermore $\mathcal{X}_0$ is accessible from $\mathcal{X}$*

$$\mathbb{P}_x(\liminf_{t\to\infty} \frac{V(\tilde{Z}_t)}{t} \geq H_*) = 1$$

for all $x \in \mathcal{X}$.

The next result is a general persistence result.

We call a point $p \in \mathcal{X}$ a *Doeblin point* provided there exists a neighborhood U of p, a bounded (positive) measure ν on $\mathcal{X}$ and some number $s > 0$ such that

$$\delta_x P_s \geq \nu$$

for all $x \in U$.

Theorem 5.4 *Suppose that the process is H-persistent. Then*

(i) *The process is stochastically persistent. In particular, for all $x \in \mathcal{X}_+$, $\mathbb{P}_x$ almost surely, every limit point of $\{\Pi_t\}$ lies in $\mathcal{P}_{inv}(\mathcal{X}_+) = \mathcal{P}_{inv} \cap \mathcal{P}(\mathcal{X}_+)$.*

(ii) *There exist* $0 < \rho < 1$ *and positive constants* $\theta > 0, K > 0, T$ *such that*

$$P_T(e^{\theta V}) \leq \rho e^{\theta V} + K;$$

(iii) *Let* $\varepsilon > 0$ *and* τ^ε *be the stopping time defined by*

$$\tau^\varepsilon = \inf\{t \geq 0 : \tilde{Z}_t \in \mathcal{X}_0^\varepsilon\}.$$

Then there exists $\varepsilon > 0$ *such that for all* $1 < b < \frac{1}{\rho}$, *there exists* $c > 0$ *such that for all* $x \in \mathcal{X}^+$

$$\mathbb{E}_x(b^\tau) \leq c(1 + e^{\theta V(x)});$$

(iv) *If, furthermore, there exists a Doeblin point* $x \in \mathcal{X}_+$ *accessible from* $\mathcal{X}_+$ *then* $\mathcal{P}_{inv}(\mathcal{X}_+)$ *reduces to a single measure* Π *and for all* $x \in \mathcal{X}_+$

$$\|\delta_x P_t - \Pi\|_{TV} \leq const.(1 + e^{\theta V(x)})e^{-\kappa t}$$

for some $\kappa > 0$.

5.2 Proofs of Theorems 3.1–3.4

In order to apply the results of the previous section we rewrite the dynamics of $Z = (X, I)$ in polar coordinates. Let $\Psi : M^* \times E \to \mathbb{R}_+^* \times S^{d-1} \times E$ be defined by $\Psi(x, i) = (\|x\|, \frac{x}{\|x\|}, i)$ and

$$\mathcal{X}_+ = \Psi(M^* \times E).$$

Whenever $X_0 \in M^*$, the process $\tilde{Z}_t = \Psi(Z_t) = (\rho_t, \Theta_t, I_t) \in \mathcal{X}_+$ satisfies the system

$$\begin{cases} \frac{d\rho_t}{dt} = \langle \Theta_t, \tilde{F}^{I_t}(\rho_t, \Theta_t)\rangle \rho_t \\ \frac{d\Theta_t}{dt} = \tilde{F}^{I_t}(\rho_t, \Theta_t) - \langle \Theta_t, \tilde{F}^{I_t}(\rho_t, \Theta_t)\rangle \Theta_t \\ \mathsf{P}(I_{t+s} = j | \mathcal{F}_t) = a_{ij}(\rho_t \Theta_t)s + o(s) \text{ for } i \neq j \text{ on } \{I_t = i\} \end{cases} \tag{18}$$

where

$$\tilde{F}^i(\rho, \theta) = \frac{F^i(\rho\theta)}{\rho}$$

for all $\rho > 0$ and $\theta \in S^{d-1}$. By C^2 continuity of F^i, the map $\tilde{F}^i$ extends to a C^1 map $\tilde{F}^i : \mathbb{R}_+ \times S^{d-1} \mapsto \mathbb{R}^d$ by setting

$$\tilde{F}^i(0, \theta) = A^i \theta.$$

Thus, using this extension, (18) extends to the state space

$$\mathcal{X} := \overline{\mathcal{X}_+} = \mathcal{X}_+ \cup \mathcal{X}_0$$

where $\mathcal{X}_0 = \{0\} \times (S^{d-1} \cap C_M) \times E$.

This induces a PDMP (still denoted $\tilde{Z}$) on $\mathcal{X}$, whose infinitesimal generator $\tilde{\mathcal{L}}$ acts on functions $f : \mathcal{X} \to \mathbb{R}$ smooths in (ρ, θ) according to

$$\tilde{\mathcal{L}}f(\rho,\theta,i) = \frac{\partial f^i}{\partial \rho}(\rho,\theta)\langle \theta, \tilde{F}^i(\rho,\theta)\rangle \rho + \langle \nabla_\theta f^i(\rho,\theta), \tilde{G}^i(\rho,\theta)\rangle + \sum_{j\in E} a_{ij}(\rho\theta)(f^j(\rho,\theta) - f^i(\rho,\theta)), \tag{19}$$

where $\tilde{G}^i(\rho,\theta) = \tilde{F}^i(\rho,\theta) - \langle \theta, \tilde{F}^i(\rho,\theta)\rangle\theta$. By [15, Proposition 2.1], $\tilde{Z}$ is Feller. Moreover by equation (18), Hypothesis 5.1 is verified. The following lemma gives V and H that fulfil Hypothesis 5.2.

Lemma 5.5 *For all $(\rho,\theta,i) \in \mathcal{X}$, set $H(\rho,\theta,i) = -\langle \tilde{F}^i(\rho,\theta), \theta\rangle$, and for $\rho \neq 0$, $V(\rho,\theta,i) = -\log(\rho)$. Then V and H satisfy Hypothesis 5.2.*

Proof The definition of $\tilde{\mathcal{L}}$ and V imply that $\tilde{\mathcal{L}}V(\rho,\theta,i) = H(\rho,\theta,i)$ for all $(\rho,\theta,i) \in \mathcal{X}_+$. For all $K \subset \mathcal{X}_+$ compact, there exists $\varepsilon > 0$ such that $\rho \geq \varepsilon$ on K. Let $\log_\varepsilon : \mathbb{R} \mapsto \mathbb{R}$ be a smooth function coinciding with log on $[\varepsilon, \infty[$. Set $V_K(\rho,\theta,i) = -\log_\varepsilon(\rho)$. Then **(a)** is satisfied, and because V_K doesn't depend on i $\Gamma(V_K) = 0$ so that **(b)** is also satisfied. **(c)** and **(d)** are clearly satisfied. Now for all $t \geq 0$, there exists operators J_u and K preserving regularity, and a Poisson Process N_t with inter arrival times U_i such that

$$P_t H = \sum_{n\geq 0} \mathbb{E}\left[\mathbb{1}_{N_t=n} J_{U_1} \circ ... \circ J_{U_n} \circ K_{t-T_n} H\right].$$

(see proof of Proposition 2.1 in [15] for details). Thus by dominated convergence and smoothness of H, **(e)** is satisfied. **QED**

Now we link the H-exponents of $\tilde{Z}$ with the extremal average growth rates of Z :

Lemma 5.6 *With the notation of the previous sections,*

$$H_* = -\Lambda^+ \quad \textit{and} \quad H^* = -\Lambda^-.$$

In particular, $\tilde{Z}$ is H-persistent if and only if $\Lambda^- > 0$ and H-nonpersistent if and only if $\Lambda^+ < 0$.

Proof On $\mathcal{X}_0$, $\tilde{Z}_t = (0, \Theta_t, J_t)$ where (Θ_t, J_t) is the process given in Section 2. Now, $\langle A^i\theta, \theta\rangle = -H(0,\theta,i)$, and the result easily follows from the definitions of $\Lambda^{+/-}, H_*, H^*$. **QED**

Thanks to these lemmas and theorems of the previous sections, we can now prove our main results.

Proof of Theorem 3.1 Here we assume $\Lambda^+ < 0$, thus by Lemma 5.6 $\tilde{Z}$ is H - nonpersistent. Theorem 5.3 **(i)** then gives exactly the first part of Theorem 3.1 because $V(\tilde{Z}_t) = -\log(\rho_t) = -\log(\|X_t\|)$ for all $x \neq 0$.

Assume furthermore that 0 is F - accessible from M. By [15, Proposition 3.14], this implies that $\{0\} \times E$ is accessible from $M \times E$ for the process Z and thus that $\mathcal{X}_0$ is accessible from $\mathcal{X}$ for the process $\tilde{Z}$. Then Theorem 5.3 **(ii)** proves the second assertion of Theorem 3.1. **QED**

To show the others theorem, we use the following lemma for which the proof is omitted.

Lemma 5.7 *The map*

$$\begin{array}{ccc} \mathcal{P}_{inv}^{\tilde{Z}}(\mathcal{X}_+) & \longrightarrow & \mathcal{P}_{inv}^{Z}(M^* \times E) \\ \Pi & \longmapsto & \Pi \circ \Psi \end{array}$$

is a bijection. Moreover, for all $(x,i) \in M^* \times E$, and all $t \geq 0$

$$\Pi_t^{x,i} = \tilde{\Pi}_t^{\Psi(x,i)} \circ \Psi.$$

Thus $\Pi_t^{x,i}$ converges almost surely to some Π if and only if $\tilde{\Pi}_t^{\Psi(x,i)}$ converges to $\Pi \circ \Psi^{-1}$.

Proof of Theorem 3.2 Here we assume $\Lambda^- > 0$, thus by Lemma 5.6 $\tilde{Z}$ is H - persistent. Then Theorem 5.4 **(i)** and Lemma 5.7 imply **(i)** of Theorem 3.2. Moreover, by Theorem 5.4 **(ii)**, we have for some positive θ, K, T

$$\tilde{P}_T(e^{\theta V}) \leq \rho e^{\theta V} + K.$$

Let $\tilde{\mu} \in \mathcal{P}_{inv}^{\tilde{Z}}(\mathcal{X}_+)$ and set $\tilde{W} = e^{\theta V}$. Then integrating the previous inequality against $\tilde{\mu}$ gives $\tilde{\mu}\tilde{W} \leq \rho\tilde{\mu}\tilde{W} + K$, thus

$$\tilde{\mu}\tilde{W} \leq \frac{K}{1-\rho}. \tag{20}$$

Now let $\mu \in \mathcal{P}_{inv}^{Z}(M^* \times E)$ and set $W(x,i) = \|x\|^{-\theta}$. Then $\mu W = (\mu \circ \Psi^{-1} \circ \Psi)W = (\mu \circ \Psi^{-1})(W \circ \Psi^{-1})$. By lemma 5.7, $\mu \circ \Psi^{-1} \in \mathcal{P}_{inv}^{\tilde{Z}}(\mathcal{X}_+)$, and because $W \circ \Psi^{-1} = \tilde{W}$, (20) proves **(ii)** of Theorem 3.2. Point **(iii)** is immediate from **(iii)** of Theorem 5.4. **QED**

Proof of Theorem 3.3 By Theorem 3.2, $\mathcal{P}_{inv}^{Z}(M^* \times E)$ is non-empty. So the weak bracket condition implies by [15, Theorem 4.5] uniqueness of Π and the absolute continuity. Moreover, for all $(x,i) \in M^* \times E$, $(\Pi_t^{x,i})_{t \geq 0}$ is tight and admits a unique limit point Π, so that $\Pi_t^{x,i}$ converges almost surely to Π. **QED**

Proof of Theorem 3.4 Assume the strong bracket condition holds at a point p that is F-accessible from M^*. Then [15, Theorem 4.2] implies that for all i, $\Psi(p,i)$ is a Doeblin point, which is accessible for the process $\tilde{Z}$ from $\mathcal{X}_+$. Thus by point **(iv)** of Theorem 5.4, for all $z = (\rho, \theta, i) \in \mathcal{X}_+$

$$\|\delta_z \tilde{P}_t - \Pi \circ \Psi^{-1}\|_{TV} \leq c(1 + \tilde{W}(z))e^{-\kappa t}.$$

Now, for all $A \in \mathcal{B}(M \times E)$ and all $(x,i) \in M^* \times E$, $\delta_{x,i}P_t(A) - \Pi(A) = \delta_{\Psi(x,i)}\tilde{P}_t(\Psi(A)) - \Pi \circ \Psi^{-1}(\Psi(A))$, so that

$$\begin{aligned} \|\delta_{x,i}P_t - \Pi\|_{TV} &= \|\delta_{\Psi(x,i)}\tilde{P}_t - \Pi \circ \Psi^{-1}\|_{TV} \\ &\leq c(1 + \tilde{W}(\Psi(x,i)))e^{-\kappa t} \\ &= c(1 + W(x,i))e^{-\kappa t}. \end{aligned}$$

Then Theorem 3.4 is proved. **QED**

6 Proof of Theorem 4.11

Before proving our convergence theorem, we first recall the definition of the Birkhoff part metric and some properties of monotone and subhomogeneous random dynamical systems given in the book of Chueshov [19]. Let D be a non-empty subset of $\{1, \dots, d\}$ and let $\mathbb{R}^d_{++,D}$ be the subset of $x \in \mathbb{R}^d_+$ such that $x_i > 0$ if $i \in D$ and $x_i = 0$ otherwise. Then $\mathbb{R}^d_{++,D}$ is called a *part*. The *Birkhoff part metric* is defined, for all $x, y \in \mathbb{R}^d_+$ by :

$$p(x, y) = \max_{i \in D} |\log(x_i) - \log(y_i)|$$

if x and y are both in the same part $\mathbb{R}^d_{++,D}$ for some D, and $p(x, y) = +\infty$ otherwise. By monotony and strong subhomogeneity of Ψ, [19, Lemma 4.2.1] ensures that Ψ is nonexpansive under the part metric on every part and strictly nonexpansive on $\mathbb{R}^d_{++}$. In other words, for all $t \geq 0$, for all $\omega \in \Omega$, for all $D \subset \{1, \dots, d\}$, for all $x, y \in \mathbb{R}^d_{++,D}$,

$$p(\Psi(t, \omega, x), \Psi(t, \omega, y)) \leq p(x, y),$$

and the inequality is strict if $D = \{1, \dots, d\}$, $x \neq y$ and $t > 0$. We would like to have a *contraction*, meaning that there exist $\alpha \in (0, 1)$ such that $p(\Psi(t, \omega, x), \Psi(t, \omega, y)) \leq \alpha p(x, y)$. The following crucial lemma states that this is true if we restrain ourselves to compact subset of $\mathbb{R}^d_{++}$.

Lemma 6.1 *Let $\varphi : \mathbb{R}^d_+ \to \mathbb{R}^d_+$ be a C^2 monotone strongly subhomogeneous map and K be a compact subset contained in $\mathbb{R}^d_{++}$. Then φ is a contraction for p on K, that is :*

$$\tau_K(\varphi) := \sup_{x, y \in K, x \neq y} \frac{p(\varphi(x), \varphi(y))}{p(x, y)} < 1.$$

Proof First note that for all $x, y \in K$, with $x \neq y$, one has $\frac{p(\varphi(x), \varphi(y))}{p(x,y)} < 1$. In particular, by continuity of p and φ, for all $\varepsilon > 0$ there exists $\alpha < 1$ such that

$$\sup_{x, y \in \Delta_\varepsilon(K)} \frac{p(\varphi(x), \varphi(y))}{p(x, y)} \leq \alpha, \tag{21}$$

where $\Delta_\varepsilon(K) = \{(x, y) \in K^2 \ : \ p(x, y) \geq \varepsilon\}$ is compact. It remains to prove that such a bound holds when x and y are close, uniformly in $x \in K$. To do so, we use the following fact: a monotone map φ is strongly sublinear if and only if, for all $x \gg 0$, $D\varphi(x)x \ll \varphi(x)$ (see e.g [19, Proposition 4.1.1] or [17, Proposition 6]). Componentwise, this means that for all i,

$$\frac{\langle \nabla \varphi_i(x), x \rangle}{\varphi_i(x)} < 1. \tag{22}$$

By Taylor expansion, for all i and all $x, y \in K$,

$$\log \varphi_i(y) - \log \varphi_i(x) = \frac{\langle \nabla \varphi_i(x), y - x \rangle}{\varphi_i(x)} + R_i(x, y)\|x - y\|^2,$$

where R_i is continuous, thus uniformly bounded on K^2 by some constant C.

Moreover, one can easily check that for all $\frac{1}{2M} \leq u \leq 2M$, one has

$$|u-1| \leq e^{|\log u|} - 1 \leq |\log u|(1 + M|\log u|).$$

Now there exists M such that for all $x, y \in K$ and k, $\frac{1}{2M} \leq y_k/x_k \leq 2M$. Thus, for all k,

$$|y_k - x_k| \leq x_k(1 + Mp(x,y))p(x,y). \tag{23}$$

For all $x, y \in \mathbb{R}^d_{++}$ and $x \neq y$, there exists i such that

$$\begin{aligned}\frac{p(\varphi(x), \varphi(y))}{p(x,y)} &= \frac{|\frac{\langle \nabla\varphi_i(x), y-x\rangle}{\varphi_i(x)} + R_i(x,y)\|x-y\|^2|}{p(x,y)} \\ &\leq \frac{|\langle \nabla\varphi_i(x), y-x\rangle|}{\varphi_i(x)p(x,y)} + |R_i(x,y)|\frac{\|x-y\|^2}{p(x,y)}\end{aligned}$$

Now by (23) and nonnegativity of $\nabla\varphi_i(x)$ (recall φ is monotone), we have for all $x, y \in K$, for all $x \neq y$,

$$\frac{p(\varphi(x), \varphi(y))}{p(x,y)} \leq \frac{\langle \nabla\varphi_i(x), x(1 + Mp(x,y))\rangle}{\varphi_i(x)} + C\frac{\|x-y\|^2}{p(x,y)}.$$

Inequality (22), continuity of φ and compactness of K implie that there exists a constant $\tau < 1$ such that, for all $x \in K$ and all i,

$$\frac{\langle \nabla\varphi_i(x), x\rangle}{\varphi_i(x)} \leq \tau,$$

and thus

$$\frac{p(\varphi(x), \varphi(y))}{p(x,y)} \leq \tau(1 + Mp(x,y)) + C\frac{\|x-y\|^2}{p(x,y)}.$$

By compactness of K, $p(x,y)$ and $\frac{\|x-y\|^2}{p(x,y)}$ converges to 0 uniformly in $x \in K$ when y converges to x. Thus, we can find $\varepsilon > 0$ such that $\tau' = \sup_{x \in K, y \in B_K(x,\varepsilon)\setminus\{x\}} \tau(1+Mp(x,y)) + C\frac{\|x-y\|^2}{p(x,y)} < 1$, where $B_K(x,\varepsilon)$ is the intersection of the ball of center x and radius ϵ with K. In other words,

$$\sup_{x,y \in \Delta^c_\varepsilon(K)} \frac{p(\varphi(x), \varphi(y))}{p(x,y)} \leq \tau'. \tag{24}$$

Combining (21) and (24) gives the result with $\tau_K(\varphi) = \max(\alpha, \tau') < 1$. **QED**

Recall that $\mathcal{Y} = [0,1]^d \setminus \{0\} \times E$ and set $d : \mathcal{Y}^2 \to [0,1]$ the distance defined by

$$d((x,i),(y,j)) = \mathbb{1}_{i \neq j} + \mathbb{1}_{i=j}(\frac{p(x,y)}{C} \wedge 1),$$

where C is a constant to be chosen later and $p(x,y)$ is the Birkhoff part metric. Define also $V : \mathcal{Y} \to \mathbb{R}_+$ with $V(x,i) = \|x\|^{-\theta}$ where θ is given in Theorem 3.2 and the function $\tilde{d} : \mathcal{Y}^2 \to \mathbb{R}_+$ by

$$\tilde{d}(z, \tilde{z}) = \sqrt{d(z,\tilde{z})(1 + V(z) + V(\tilde{z}))}.$$

As already mentioned, Theorem 4.11 is a consequence of the weak form of Harris' theorem due to Hairer, Mattingly and Scheutzow [27, Theorem 4.8 and remark 4.10]. More precisely, it states that point **(i)** of Theorem 4.11 holds, provided the three following assumptions are verified (here we let P_t denoted P^Z_t) :

A1 V is a *Lyapunov function* for P_t, that is there exists $C_V, \gamma, K_V, t_0 > 0$ such that for all $t \geq t_0$, for all $z \in \mathcal{X}$,

$$P_t V(z) \leq C_V e^{-\gamma t} V(x) + K_V;$$

A2 There exists $t^* > t_* > 0$ such that for all $t \in [t_*, t^*]$, the level set $A_V = \{z \in \mathcal{X} \,:\, V(x) \leq 4K_V\}$ are *d-small* for P_t, meaning that there exists $\varepsilon > 0$ such that for all $z, \tilde{z} \in A_V$,

$$\mathcal{W}_d(\delta_z P_t, \delta_{\tilde{z}} P_t) \leq 1 - \varepsilon;$$

A3 For all $t \in [t_*, t^*]$, P_t is *contracting* on A_V, meaning that there exists $\alpha \in (0,1)$ such that for all $z, \tilde{z} \in A_V$ with $d(z, \tilde{z}) < 1$,

$$\mathcal{W}_d(\delta_z P_t, \delta_{\tilde{z}} P_t) \leq \alpha d(z, \tilde{z}).$$

Moreover, P_t is nonexpansive on $\mathcal{X}$, that is for all $z, \tilde{z} \in \mathcal{X}$,

$$\mathcal{W}_d(\delta_z P_t, \delta_{\tilde{z}} P_t) \leq d(z, \tilde{z}).$$

Remark 6.2 In [27, Theorem 4.8], the hypothesis **A1** and **A3** are a little bit stronger : **A1** should holds for every $t \geq 0$, and the contraction in **A3** should holds on the whole space $\mathcal{X}$ for $d(z, \tilde{z}) < 1$. However, a quick look at the proof given in [27] shows that it is enough to have the Lyapunov function for t large, and that when $z, \tilde{z}$ are such that $1 + V(z) + V(\tilde{z}) \geq 4K_V$, the proof "Far from the origin" is true independently from the fact that $d(z, \tilde{z}) < 1$ or $d(z, \tilde{z}) \geq 1$

□

To prove Theorem 4.11 it is thus sufficient to show that **A1** to **A3** are satisfied. For **A1**, it is a consequence of a stochastic persistence lemma. For **A2**, we show that a good choice of the constant C appearing in the definition of d is sufficient to have the small set. Finally, **A3** is a consequence of the contracting properties of $\Psi(t, \omega)$.

Proof of Theorem 4.11

A1 We have the following lemma :

Lemma 6.3 *For $0 < \alpha < \lambda_1$, there exists $T > 0$, $\varepsilon > 0$ and $C > 0$ such that, for all $t \in [T, 3T/2]$, for all $z \in \mathcal{Y}_0^\varepsilon$,*

$$P_t V(z) \leq e^{\theta t (\frac{t}{T} - 1)\alpha} V(z),$$

where $\theta = \frac{\alpha}{CT}$, $\mathcal{Y}_0^\varepsilon = \{(x, i) \in \mathcal{Y} \,:\, \|x\| < \varepsilon\}$ and $V(x, i) = \|x\|^{-\theta}$.

Proof Follows the lines of the proof given in [16, Lemma 3.5]. **QED**

In particular, putting $\gamma = \frac{\theta\alpha}{4}$, then for all $t \in [T, 3T/2]$, for all $z \in \mathcal{Y}_0^\varepsilon$,

$$P_t V(z) \leq e^{\gamma t} V(z).$$

Now by Feller continuity of P_t and compactness of $[T, 3T/2] \times \mathcal{Y} \setminus \mathcal{Y}_0^\varepsilon$

$$\tilde{C} = \sup_{(t,z) \in [T, 3T/2] \times \mathcal{Y} \setminus \mathcal{Y}_\varepsilon} P_t V(z) - V(z) < \infty,$$

and, for all $t \in [T, 3T/2]$ and all $z \in \mathcal{Y}$,

$$P_t V(z) \leq e^{\gamma t} V(z) + \tilde{C}.$$

If $t \geq 2T$, then there exists $s \in [T, 3T/2]$ and $n \geq 1$ such that $t = ns$. Thus

$$P_t V(z) = P_{ns} V(z) \leq e^{\gamma ns} V(z) + \sum_{k=0}^{n-1} e^{\gamma ks} \tilde{C},$$

proving **A1** with $t_0 = 2T$ and $K_V = \frac{1}{1-e^{-\gamma T}} \tilde{C}$.

A2 Set $M_V = \{x \in [0,1]^d \setminus \{0\} \, : \, \|x\|^{-\theta} \leq 4K_V\}$. We first prove that for all $t^* > t_* > 0$, there exists a compact set contained in $\mathbb{R}^d_{++}$ such that for all $t \in [t_*, t^*]$, and all $\omega \in \Omega$, $\Psi(t, \omega, M_V)$ is included in this compact. For this, let S_{M_V} denotes the set of all the solutions of the differential inclusion

$$\begin{cases} \dot{\eta}(t) \subset \mathsf{co}(\tilde{\mathrm{F}})(\eta(t)) \\ \eta(0) = x, \end{cases}$$

with $x \in M_V$. Then because M_V is compact, S_{M_V} is a non avoid compact subset of $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ (see e.g Aubin and Cellina [4, Section 2.2 Theorem 1]). This implies that $\mathbf{\Psi}_{[t_*,t^*]}(M_V) = \{\eta_t \, : \, t \in [t_*, t^*], \eta \in S_{M_V}\}$ is a compact set of $[0,1]^d$. Moreover, by strong monotony of η_t, $\mathbf{\Psi}_{[t_*,t^*]}(M_V)$ is included in $(0,1]^d$ and for all $t \in [t_*, t^*]$, $\omega \in \Omega$, $\Psi(t, \omega, M_v) \subset \mathbf{\Psi}_{[t_*,t^*]}(M_V)$. Now by compactness of $\mathbf{\Psi}_{[t_*,t^*]}(M_V)$ and continuity of p, there exist $K > 0$ such that for all $t \in [t_*, t^*]$,

$$\sup_{x,y \in M_v; \omega, \omega' \in \Omega} p(\Psi(t, \omega, x), \Psi(t, \omega', y)) \leq \sup_{a,b \in \mathbf{\Psi}_{[t_*,t^*]}(M_V)} p(a,b) = K. \tag{25}$$

To prove **A2**, for any $(z, \tilde{z}) = ((x,i),(y,j)) \in \mathcal{Y}^2$, we consider the coupling $(Z_t, \tilde{Z}_t) = ((X_t, I_t), (Y_t, J_t))$ of $\delta_z P_t$ and $\delta_{\tilde{z}} P_t$ construct as follows. If $i = j$, then $I_t = J_t$ for all $t \geq 0$. If $i \neq j$, then I_t and J_t evolves independently until the first meeting time T and then are stick together for ever. In other words,

$$\mathbb{P}_{i,j}(I_t \neq J_t) = \mathbb{P}_{i,j}(T > t).$$

This is the coupling considered in [13]. As stated in [13, Lemma 2.1], we easily control the above probability : there exists $\rho > 0$ such that for all $i, j \in E$ and all $t \geq 0$,

$$\mathbb{P}_{i,j}(I_t \neq J_t) = \mathbb{P}_{i,j}(T > t) \leq e^{-\rho t}.$$

Let $(z, \tilde{z}) = ((x,i),(y,j)) \in A_V^2$ and $t \in [t_*, t^*]$. Then

$$\begin{aligned} \mathcal{W}_d(\delta_z P_t, \delta_{\tilde{z}} P_t) &\leq \mathbb{E}_{(z,\tilde{z})}(d(Z_t, \tilde{Z}_t)) \\ &\leq \mathbb{P}_{i,j}(I_t \neq J_t) + \mathbb{E}_{(z,\tilde{z})}(\frac{p(X_t, Y_t)}{C}) \\ &\leq e^{-\rho t} + \frac{K}{C}, \end{aligned}$$

where the last inequality comes from (25). Thus, choosing $C = \frac{K}{1-2e^{-\rho t_*}}$, one has

$$\mathcal{W}_d(\delta_z P_t, \delta_{\tilde{z}} P_t) \leq 1 + e^{-\rho t} - 2e^{-\rho t_*} \leq 1 - e^{-\rho t_*},$$

proving **A2** with $\varepsilon = e^{-\rho t_*}$.

A3 We first prove that P_t is nonexpansive on $\mathcal{Y}$. Is suffices to show the result for $(z,\tilde{z})$ such that $d(z,\tilde{z})<1$, the bound being trivial otherwise. In particular, $i=j$ where $z=(x,i)$ and $\tilde{z}=(y,j)$, and $d(z,\tilde{z})=\frac{p(x,y)}{C}<1$, which implies that x and y are in the same part. We consider the same coupling $(Z_t,\tilde{Z}_t)$ as above. Then because $i=j$, $I_t=J_t$ and thus $X_t=\Psi(t,\omega,x)$ and $Y_t=\Psi(t,\omega,y)$, and so by nonexpansivity of $\Psi(t,\omega)$ on every part, one has $p(\Psi(t,\omega,x),\Psi(t,\omega,y))\le p(x,y)$, which gives the result for P_t.

Now we prove that P_t is a contraction on A_V. Let $t\in[t_*,t^*]$ and $(z,\tilde{z})\in A_V^2$ such that $d(z,\tilde{z})<1$. In addition with the consequences cited above, this also implies that $x,y\in M_V$. Choose $0<t_0<t_*$, then one has

$$\begin{aligned}p(\Psi(t,\omega,x),\Psi(t,\omega,y)) &= p(\Psi(t-t_0+t_0,\omega,x),\Psi(t-t_0+t_0,\omega,y))\\ &\le p(\Psi(t-t_0,\Theta_{t_0}\omega)\Psi(t_0,\omega,x),\Psi(t-t_0,\Theta_{t_0}\omega)\Psi(t_0,\omega,y))\\ &\le \tau_{\mathbf{\Psi}_{t_0}(M_V)}(\Psi(t-t_0,\Theta_{t_0}\omega))p(\Psi(t_0,\omega,x),\Psi(t_0,\omega,y))\\ &\le \tau_{\mathbf{\Psi}_{t_0}(M_V)}(\Psi(t-t_0,\Theta_{t_0}\omega))p(x,y),\end{aligned}$$

where $\tau_{\mathbf{\Psi}_{t_0}(M_V)}(\Psi(t-t_0,\Theta_{t_0}\omega))<1$ is the contraction constant given by Lemma 6.1 on the compact $\mathbf{\Psi}_{t_0}(M_V)\subset\mathbb{R}^d_{++}$. Because $\tau_{\mathbf{\Psi}_{t_0}(M_V)}(\Psi(t-t_0,\Theta_{t_0}\omega))<1$ for every ω, then

$$\alpha=\max_i \mathbb{E}_i[\tau_{\mathbf{\Psi}_{t_0}(M_V)}(\Psi(t-t_0,\Theta_{t_0}\omega))]<1,$$

and

$$\mathcal{W}_d(\delta_z P_t,\delta_{\tilde{z}}P_t)\le \mathbb{E}_{(x,i),(y,j)}(\frac{p(\Psi(t,\omega,x),\Psi(t,\omega,y))}{C}\le \alpha\frac{p(x,y)}{C}=\alpha d(z,\tilde{z}),$$

proving **A3** and the **(i)** of the theorem.

Because $\lambda_1>0$, Theorem 3.2 insures existence of an invariant measure for P_t on $\mathcal{Y}$. The uniqueness of the invariant measure and thus point **(ii)** follows immediately from point **(i)**. **QED**

7 Appendix

7.1 Proof of Proposition 2.11

Recall (see section 4) that $\mathbb{R}^d_{++}$ denotes the interior of $\mathbb{R}^d_+$, (i.e the cone of positive vectors). Set $S^{d-1}_+=S^{d-1}\cap\mathbb{R}^d_+$ and $S^{d-1}_{++}=S^{d-1}\cap\mathbb{R}^d_{++}$. The principal tool is the *projective* or *Hilbert metric* d_H on $\mathbb{R}^d_{++}$ (see Seneta [43]) defined by

$$d_H(x,y)=\log\frac{\max_{1\le i\le d}x_i/y_i}{\min_{1\le i\le d}x_i/y_i}.$$

Note that

$$d_H(\frac{x}{\|x\|},\frac{y}{\|y\|})=d_H(x,y) \tag{26}$$

so that d_H is not a distance on $\mathbb{R}^d_{++}$. However its restriction to S^{d-1}_{++} is. Furthermore, for all $x,\, y\in S^{d-1}_{++}$,

$$\|x-y\|\le \mathrm{e}^{d_H(x,y)}-1. \tag{27}$$

Let $\mathcal{M}_+$ denote the set of $d \times d$ Metzler matrices having **positive** diagonal entries, and let $\mathcal{M}_{++} \subset \mathcal{M}_+$ denote the set of matrices having positive entries. By a theorem of Garret Birkhoff, there exists a continuous map $\tau : \mathcal{M}_{++} \mapsto]0,1[$ such that for all $T \in \mathcal{M}_{++}$, and all $x, y \in \mathbb{R}^d_{++}$

$$d_H(Tx, Ty) \leq \tau[T] d_H(x,y) \tag{28}$$

The number $\tau[T]$ is usually called the *Birkhoff's contraction coefficient* of T, and is given by an explicit formulae (see e.g [43], Section 3.4) which is unneeded here.

We extend τ to a measurable map $\tau : \mathcal{M}_+ \mapsto]0,1]$ by setting $\tau[T] = 1$ for all $T \in \mathcal{M}_+ \setminus \mathcal{M}_{++}$. By density of $\mathcal{M}_{++}$ in $\mathcal{M}_+$ and continuity of d_H on $\mathbb{R}^d_{++}$ it is easy to see that (28) extends to $\mathcal{M}_+$.

For each $\omega \in \Omega$, the map $t \mapsto \varphi(t,\omega)$ is solution to the matrix valued differential equation

$$\forall t \geq 0,\ \frac{dM}{dt} = A^{\omega_t} M, M_0 = I_d. \tag{29}$$

Thus,

$$\varphi(t,\omega) \in \mathcal{M}_+$$

for all $t \geq 0$. Indeed, for all $i \in E$ and $r > 0$ large enough $A^i + rI_d \in \mathcal{M}_+$, so that $e^{tA^i} = e^{-rt} e^{t(A^i + rI_d)} \in \mathcal{M}_+$.

We claim that there exists a Borel set $\tilde{\Omega} \subset \Omega$ with $\mathbb{P}^J_i(\tilde{\Omega}) = 1$ for all $i \in E$, and such that for all $\omega \in \tilde{\Omega}$:

(i) $\exists n \in \mathbb{N}\, \varphi(n,\omega) \in \mathcal{M}_{++}$;

(ii) $\forall n \in \mathbb{N}\ \limsup_{t\to\infty} \dfrac{\log \tau[\varphi(t, \mathbf{\Theta}_n(\omega))]}{t} < 0.$

Before proving these assertions let us show how they imply the result to be proved. For all $\omega \in \tilde{\Omega}$ and n given by (i),

$$\varphi(t+n,\omega) = \varphi(t, \mathbf{\Theta}_n(\omega))\varphi(n,\omega) \in \mathcal{M}_{++}$$

as the product of an element of $\mathcal{M}_+$ with an element of $\mathcal{M}_{++}$. Thus, by (ii), for all $\omega \in \tilde{\Omega}$ and $x, y \in \mathbb{R}^d_+ \setminus \{0\}$

$$\limsup_{t\to\infty} \frac{1}{t} \log d_H(\varphi(t+n,\omega)x, \varphi(t+n,\omega)y) < 0. \tag{30}$$

For $x \in S^{d-1}_+$ set

$$\Phi(t,\omega)x = \frac{\varphi(t,\omega)x}{\|\varphi(t,\omega)x\|}.$$

Let $f : S^{d-1}_+ \times E \to \mathbb{R}$ be a continuous map. It follows from (30), (26), (27) and the continuity of f that

$$|f(\Phi(t,\omega)x, \omega_t) - f(\Phi(t,\omega)y, \omega_t)| \to 0$$

for all $x, y \in S^{d-1}_+$ and $\omega \in \tilde{\Omega}$. Moreover,

$$P^{(\Theta,J)}_t f(x,i) = \mathbb{E}^J_i(f(\Phi(t,\omega)x, \omega_t)),$$

and thus

$$\lim_{t\to\infty} P_t^{(\Theta,J)} f(x,i) - P_t^{(\Theta,J)} f(y,i) = \lim_{t\to\infty} \mathbb{E}_i^J (f(\Phi(t,\omega)x,\omega_t) - f(\Phi(t,\omega)y,\omega_t)) = 0$$

by dominated convergence. Now take $\mu,\, \nu \in \mathcal{P}_{inv}^{(\Theta,J)}$. Then one has

$$\lim_{t\to\infty} \sum_i p_i \int_{(S_+^{d-1})^2} \left(P_t^{(\Theta,J)} f(x,i) - P_t^{(\Theta,J)} f(y,i) \right) \mu(\mathrm{d}x|i)\nu(\mathrm{d}y|i) = 0, \tag{31}$$

where $\mu(\cdot|i) = \mu^i(\cdot)/p_i$. But by invariance of μ and ν, the left-hand side of (31) equals $\mu f - \nu f$ for all t, giving $\mu f = \nu f$ for all continuous f. This proves unique ergodicity of (Θ, J).

We now pass to the proofs of assertions (i) and (ii) claimed above.

Irreducibility of $\overline{A}$ implies that $e^{\overline{A}} \in \mathcal{M}_{++}$. Let $\mathcal{U} \subset \mathcal{M}_{++}$ be a compact neighborhood of $e^{\overline{A}}$. Since $\overline{A}.M \in \mathsf{co}(A^i)(M)$, it follows from the Support Theorem ([15, Theorem 3.4]), applied to the PDMP (29), that for all $i \in E$

$$\mathbb{P}_i^J\{\omega \in \Omega : \ \varphi(1,\omega) \in \mathcal{U}\} > 0.$$

Thus, by the Markov property or the conditional version of the Borel Cantelli Lemma, for $\mathbb{P}_i^J$ almost all ω, $\varphi(1, \mathbf{\Theta}_n(\omega)) \in \mathcal{U}$ for infinitely many n, and consequently, for n large enough

$$\varphi(n,\omega) = \varphi(1, \mathbf{\Theta}_{n-1}\omega) \dots \varphi(1,\omega) \in \mathcal{M}_{++}.$$

This proves assertion (i). By the cocycle property and Birkhoff ergodic theorem, for $\mathbb{P}_p^J$ (hence $\mathbb{P}_i^J$) almost all ω

$$\limsup_{t\to\infty} \frac{1}{t} \log(\tau[\varphi(t,\omega)]) \leq \limsup_{n\to\infty} \frac{1}{n} \log(\tau[\varphi(n,\omega)]) \leq \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log\left(\tau[\varphi(1, \mathbf{\Theta}_{k-1}(\omega))]\right)$$

$$= \mathsf{E}_p^J(\log(\tau[\varphi(1,\omega)])) \leq \sup_{M\in\mathcal{U}} \log(\tau[M])\mathbb{P}_p^J(\omega \in \Omega \ : \varphi(1,\omega) \in \mathcal{U}) < 0.$$

Replacing ω par $\mathbf{\Theta_n}(\omega)$ proves assertion (ii). **QED**

7.2 Proof of Lemma 2.12

Before proving Lemma 2.12, we prove the following lemma, which is a consequence of results from Freidlin and Wentzell [24].

Lemma 7.1 *Assume the switching rates are constant and depend on a small parameter ε : $a_{i,j}^\varepsilon = a_{i,j}/\varepsilon$ where $(a_{i,j})$ is an irreducible matrix with invariant probability p. Denote by $(X^\varepsilon, J^\varepsilon)$ the PDMP associated with $a_{i,j}^\varepsilon$ given by (2). Let Ψ denote the flow induced by the average vector field $F^p := \sum_i p_i F^i$ Then for all $\delta > 0$ and all $T > 0$,*

$$\lim_{\varepsilon\to 0} \mathbb{P}_{(x,i)} \left(\max_{0\leq t\leq T} |X_t^\varepsilon - \Psi_t(x)| > \delta \right) = 0, \tag{32}$$

uniformly in $(x,i) \in M \times E$.

Proof According to [24, Chapter 2 Theorem 1.3], it suffices to show that for all $\delta > 0$ and all $T > 0$,

$$\lim_{\varepsilon \to 0} \mathbb{P}_i^J \left(\left| \int_{t_0}^{t_0+T} (F^{J_t^\varepsilon}(x) - F^p(x)) \mathrm{d}t \right| > \delta \right) = 0, \tag{33}$$

uniformly in $t_0 > 0$ and $(x,i) \in M \times E$. Note that

$$\begin{aligned}
\left| \int_{t_0}^{t_0+T} (F^{J_t^\varepsilon}(x) - F^p(x)) \mathrm{d}t \right| &= \left| \int_{t_0}^{t_0+T} (\sum_j F^j(x) \mathbb{1}_{J_t^\varepsilon = j} - \sum_j p_j F^j(x)) \mathrm{d}t \right| \\
&\leq \sum_j \|F^j\|_\infty \left| \int_{t_0}^{t_0+T} (\mathbb{1}_{J_t^\varepsilon = j} - p_j) \mathrm{d}t \right|,
\end{aligned}$$

so (33) is proven if we show that $\int_{t_0}^{t_0+T} \mathbb{1}_{J_t^\varepsilon = j} \mathrm{d}t$ converges in probability to $p_j T$ uniformly in $t_0 > 0$. By Fubini's Theorem and invariance of p, $\mathbb{E}_p^J \left(\int_{t_0}^{t_0+T} \mathbb{1}_{J_t^\varepsilon = j} \mathrm{d}t \right) = p_j T$, so Bienaymé - Tschebischev inequality gives

$$\mathbb{P}_i^J \left(\left| \int_{t_0}^{t_0+T} (\mathbb{1}_{J_t^\varepsilon = j} - p_j) \mathrm{d}t \right| > \delta \right) \leq \frac{V_p^J (\int_{t_0}^{t_0+T} (\mathbb{1}_{J_t^\varepsilon = j} \mathrm{d}t)}{\delta},$$

where V_p^J is the variance associated to $\mathbb{E}_p^J$. Hence we can conclude if $\mathbb{E}_p^J \left[\left(\int_{t_0}^{t_0+T} \mathbb{1}_{J_t^\varepsilon = j} \mathrm{d}t \right)^2 \right]$ converges to $(p_j T)^2$ uniformly in $t_0 > 0$.

Denote by Q the intensity matrix of J^1, then for all $\varepsilon > 0$, the intensity matrix of J^ε is Q/ε and for all $i, j \in E$ and $t \geq 0$,

$$\mathbb{P}_i(J_t^\varepsilon = j) = \left(e^{\frac{t}{\varepsilon} Q} \right)_{i,j}.$$

By ergodicity of J_t^ε, the above quantity goes to p_j when $t \to \infty$ so also for every fixed t when ε goes to 0. Now we have

$$\begin{aligned}
E_p^J \left[\left(\int_{t_0}^{t_0+T} \mathbb{1}_{J_t^\varepsilon = j} \mathrm{d}t \right)^2 \right] &= 2 \int_{t_0}^{t_0+T} \int_{t_0}^{t} \mathbb{P}_p \left(J_u^\varepsilon = j; J_t^\varepsilon = j \right) \mathrm{d}u \mathrm{d}t \\
&= 2 \int_{t_0}^{t_0+T} \int_{t_0}^{t} \mathbb{P}_j \left(J_{t-u}^\varepsilon = j \right) p_j \mathrm{d}u \mathrm{d}t \\
&= 2 \int_{t_0}^{t_0+T} \int_{t_0}^{t} \left(e^{\frac{t-u}{\varepsilon} Q} \right)_{j,j} p_j \mathrm{d}u \mathrm{d}t,
\end{aligned}$$

where the second inequality resulted from the Markov property. Now because for all t_0, $t - u \in [0, T]$, $\left(e^{\frac{t-u}{\varepsilon} Q} \right)_{j,j}$ converges almost everywhere to p_j and thus the lemma is proven by dominated convergence. **QED**

With the notation of the preceding lemma, let

$$\mu^\varepsilon \in \mathcal{P}_{inv}^{(X^\varepsilon, J^\varepsilon)},\ \nu^\varepsilon = \sum_i \mu^{i,\varepsilon}.$$

The proof of the next lemma is similar to the proof of [8, Corollary 3.2].

Lemma 7.2 *Let ν a limit point of (ν^ε) when $\varepsilon \to 0$. Then ν is an invariant measure of F^p.*

Proof For notational convenience, we assume that ν^ε converges to ν.

Let $g : M \to \mathbb{R}$ be a continuous map, then for all $t > 0$ and all $\varepsilon > 0$,

$$\begin{aligned}
\left|\int g(\Psi_t)d\nu - \int gd\nu\right| &\le \left|\int g(\Psi_t)d\nu - \int gd\nu^\varepsilon\right| + \left|\int gd\nu^\varepsilon - \int gd\nu\right| \\
&\le \left|\int g(\Psi_t)d\nu - \int g(\Psi_t)d\nu^\varepsilon\right| + \left|\int g(\Psi_t)d\nu^\varepsilon - \int \mathbb{E}(g(\Theta_t^\varepsilon))d\nu^\varepsilon\right| \\
&\quad + \left|\int gd\nu^\varepsilon - \int gd\nu\right|,
\end{aligned}$$

where we have use invariance of ν and ν^ε. The first and the last term of the right hand side converge to 0 by definition of ν, and the second one also converges to 0 by Lemma 7.1. **QED**

Now let μ be a limit point of (μ^ε). For notational convenience, we assume that μ^ε converges to μ. We prove that $\mu = \nu \otimes p$, which implies Lemma 2.12. For every continuous $f : M \times E \to \mathbb{R}$, every $t \ge 0$ and $\varepsilon > 0$, one has

$$\begin{aligned}
\mu^\varepsilon f - \mu f &= \int_{M\times E} \mathbb{E}_{(x,i)}\left(f_{J_t^\varepsilon}(X_t^\varepsilon)\right) d\mu^\varepsilon(x,i) - \sum_j p_j \int_M f_j(\Psi_t(x))d\nu(x) \\
&= \int_{M\times E} \mathbb{E}_{(x,i)}\left(f_{J_t^\varepsilon}(X_t^\varepsilon)\right) d\mu^\varepsilon(x,i) - \int_{M\times E} \mathbb{E}_{(x,i)}\left(f_{J_t^\varepsilon}(\Psi_t)\right) d\mu^\varepsilon(x,i) \\
&\quad + \int_{M\times E} \mathbb{E}_{(x,i)}\left(f_{J_t^\varepsilon}(\Psi_t)\right) d\mu^\varepsilon(x,i) - \sum_j p_j \int_{M\times E} f_j(\Psi_t(x))d\mu^\varepsilon(x,i) \\
&\quad + \sum_j p_j \int_{M\times E} f_j(\Psi_t(x))d\mu^\varepsilon(x,i) - \sum_j p_j \int_M f_j(\Psi_t(x))d\nu(x) \\
&= A + B + C.
\end{aligned}$$

We have

$$\sup_{(x,i)\in M\times E} \mathbb{E}_{(x,i)}\left|f_{J_t^\varepsilon}(X_t^\varepsilon) - f_{J_t^\varepsilon}(\Psi_t)\right)| \le \max_j \sup_{(x,i)\in M\times E} \mathbb{E}_{(x,i)}\left(f_j(X_t^\varepsilon) - f_j(\Psi_t)\right),$$

where the right hand side converges to 0 when ε goes to 0 thanks to Lemma 7.1, so A converges to 0. Next,

$$|B| \le \sum_j \int_{M\times E} |\mathbb{P}_i(J_t^\varepsilon = j) - p_j|\,|f_j(\Psi_t(x))|d\mu^\varepsilon(x,i),$$

because $\mathbb{E}_{(x,i)}\left(f_{J_t^\varepsilon}(\Psi_t)\right) = \sum_j \mathbb{P}_i(J_t^\varepsilon = j) f_j(\Psi_t(x))$. Thus B converges to 0 because $|\mathbb{P}_i(J_t^\varepsilon = j) - p_j|$ converges to 0 uniformly in i and j. Finally, by definition of ν^ε

$$C = \int_M \sum_j p_j f_j(\Psi_t(x)) d\mu^{1,\varepsilon}(x,i) - \int_M \sum_j p_j f_j(\Psi_t(x)) d\nu(x),$$

proving that C converges to 0 by definition of ν and thus the Lemma.

QED

Acknowledgments

This work was supported by the SNF grant 2000020−149871/1. ES thanks Carl-Erik Gauthier for many discussions on this subject.

References

[1] M. Ait Rami, V. S. Bokharaie, O. Mason, and F. R. Wirth, *Stability criteria for SIS epidemiological models under switching policies*, Discrete Contin. Dyn. Syst. Ser. B **19** (2014), no. 9, 2865–2887. MR 3261402

[2] L. Arnold, *Random dynamical systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998. MR 1723992

[3] L. Arnold, V. Gundlach, and L. Demetrius, *Evolutionary formalism for products of positive random matrices*, Ann. Appl. Probab. **4** (1994), no. 3, 859–901. MR 1284989

[4] J-P. Aubin and A. Cellina, *Differential inclusions: set-valued maps and viability theory*, vol. 264, Springer Science & Business Media, 2012.

[5] Y. Bakhtin and T. Hurth, *Invariant densities for dynamical systems with random switching*, Nonlinearity (2012), no. 10, 2937–2952.

[6] Y. Bakhtin, T. Hurth, and J.C. Mattingly, *Regularity of invariant densities for 1d-systems with random switching*, Nonlinearity **28** (2015), 3755–3787.

[7] P. Baxendale, *Invariant measures for nonlinear stochastic differential equations*, Lyapunov exponents (Oberwolfach, 1990), Lecture Notes in Math., vol. 1486, Springer, Berlin, 1991, pp. 123–140. MR 1178952

[8] M. Benaïm, *Recursive algorithms, urn processes and chaining number of chain recurrent sets*, Ergodic Theory Dynam. Systems **18** (1998), no. 1, 53–87. MR 1609499

[9] M. Benaïm, *Stochastic persistence*, Preprint, 2014.

[10] M. Benaïm, F. Colonius, and R. Lettau, *Supports of Invariant Measures for Piecewise Deterministic Markov Processes*, ArXiv e-prints (2016).

[11] M. Benaïm and M. W. Hirsch, *Differential and stochastic epidemic models*, Differential equations with applications to biology (Halifax, NS, 1997), Fields Inst. Commun., vol. 21, Amer. Math. Soc., Providence, RI, 1999, pp. 31–44. MR 1662601

[12] 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.

[13] 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.

[14] ______, *On the stability of planar randomly switched systems*, Annals of Applied Probabilities (2014), no. 1, 292–311.

[15] ______, *Qualitative properties of certain piecewise deterministic markov processes*, Annales de l'IHP **51** (2015), no. 3, 1040 – 1075.

[16] 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.

[17] M. Benaïm and S. Schreiber, *Persistence of structured populations in random environments*, Theoretical Population Biology **76** (2009), 19–34.

[18] P. L. Chesson, *The stabilizing effect of a random environment*, Journal of Mathematical Biology **15** (1982), 1–36.

[19] I. Chueshov, *Monotone random systems theory and applications*, vol. 1779, Springer Science & Business Media, 2002.

[20] B. Cloez and M. Hairer, *Exponential ergodicity for markov processes with random switching*, Bernoulli (2015), no. 1, 505–536.

[21] F. Colonius and G. Mazanti, *Lyapunov exponents for random continuous-time switched systems and stabilizability*, ArXiv e-prints (2015).

[22] 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

[23] L. Fainshil, M. Margaliot, and P. Chigansky, *On the stability of positive linear switched systems under arbitrary switching laws*, IEEE Trans. Automat. Control **54** (2009), no. 4, 897–899. MR 2514831

[24] M. Freidlin and A. Wentzell, *Random perturbations of dynamical systems*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 260, Springer, Heidelberg, 2012, Translated from the 1979 Russian original by Joseph Szücs. MR 2953753

[25] 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.

[26] L. Gurvits, R. Shorten, and O. Mason, *On the stability of switched positive linear systems*, IEEE Transactions on Automatic Control **52** (2007), no. 6, 1099–1103.

[27] M. Hairer, J. C. Mattingly, and M. Scheutzow, *Asymptotic coupling and a general form of harris theorem with applications to stochastic delay equations*, Probability Theory and Related Fields **149** (2011), no. 1-2, 223–259.

[28] R. Z. Has'minskiĭ, *Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations*, Teor. Verojatnost. i Primenen. **5** (1960), 196–214. MR 0133871

[29] A. Hening, D. H. Nguyen, and G. Yin, *Stochastic population growth in spatially heterogeneous environments: The density-dependent case*, ArXiv e-prints (2016).

[30] M. W. Hirsch, *Positive equilibria and convergence in subhomogeneous monotone dynamics*, Comparison methods and stability theory (Waterloo, ON, 1993), Lecture Notes in Pure and Appl. Math., vol. 162, Dekker, New York, 1994, pp. 169–188. MR 1291618

[31] J. Hofbauer and S. Schreiber, *To persist or not to persist* ?, Nonlinearity **17** (2004), 1393–1406.

[32] G. Lagasquie, *A note on simple randomly switched linear systems*, arXiv preprint arXiv:1612.01861 (2016).

[33] A. Lajmanovich and J. Yorke, *A deterministic model for gonorrhea in a nonhomogeneous population*, Math. Biosci. **28** (1976), no. 3/4, 221–236. MR 0403726

[34] S. D. Lawley, J. C. Mattingly, and M. C. Reed, *Sensitivity to switching rates in stochastically switched odes*, Commun Math Sci. (2014), no. 7, 1343–1352.

[35] F. Malrieu, *Some simple but challenging Markov processes*, Ann. Fac. Sci. Toulouse Math. (6) **24** (2015), no. 4, 857–883. MR 3434260

[36] F. Malrieu and T. Hoa Phu, *Lotka-Volterra with randomly fluctuating environments: a full description*, ArXiv e-prints (2016).

[37] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability, Second Edition.*, Cambridge University Press, 2009.

[38] J. Mierczyński, *Lower estimates of top Lyapunov exponent for cooperative random systems of linear ODEs*, Proc. Amer. Math. Soc. **143** (2015), no. 3, 1127–1135. MR 3293728

[39] G. Roth and S. Schreiber, *Persistence in fluctuating environments for interacting structured populations*, Journal of Mathematical Biology **69** (2014), no. 5, 1267–1317.

[40] S. Schreiber, *Criteria for c^r robust permanence*, J Differential Equations (2000), 400–426.

[41] ———, *Persistence for stochastic difference equations: A mini review*, Journal of Difference Equations and Applications **18** (2012), 1381–1403.

[42] S. Schreiber, M. Benaïm, and Atchadé K. A. S., *Persistence in fluctuating environments*, Journal of Mathematical Biology **62** (2011), 655–683.

[43] E. Seneta, *Non-negative matrices and Markov chains*, Springer Series in Statistics, Springer, New York, 2006, Revised reprint of the second (1981) edition [Springer-Verlag, New York; MR0719544]. MR 2209438