Persistance de Processus de Markov Déterministes par Morceaux

THÈSE

présentée à la Faculté des Sciences pour obtenir le grade de docteur ès Sciences en Mathématiques par

Edouard Antoine Strickler

Soutenue le 21 Mars 2019 devant le jury composé de

Michel BENAÏM	Université de Neuchâtel, CH	Directeur de thèse
Nicolas CHAMPAGNAT	INRIA Nancy, FR	Examinateur externe
Florent MALRIEU	Université de Tours, FR	Examinateur externe
Denis VILLEMONAIS	Université de Lorraine, FR	Examinateur externe
Felix SCHLENK	Université de Neuchâtel, CH	Examinateur interne

Au vu des rapports de

Florent MALRIEU	Université de Tours, FR
Sebastian SCHREIBER	University of California, Davis, USA
George YIN	Wayne State University, USA

Institut de mathématiques, Université de Neuchâtel
Rue Emile Argand 11, CH-2000 Neuchâtel

UNIVERSITÉ DE NEUCHÂTEL

Faculté des Sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel – Suisse
Tél : + 41 (0)32 718 21 00
E-mail : secretariat.sciences@unine.ch

IMPRIMATUR POUR THÈSE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Edouard STRICKLER

Titre:

"Persistance de Processus de Markov Déterministes par Morceaux"

sur le rapport des membres du jury composé comme suit:

- Prof. Michel Benaïm, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Felix Schlenk, Université de Neuchâtel, Suisse
- Prof. Laurent Malrieu, Université de Tours, France
- Dr Denis Villemonais, Université de Lorraine, France
- Dr Nicolas Champagnat, INRIA, Nancy, France
- Prof. Sebastian Schreiber, University of California, USA
- Prof. George Yin, Wayne State University, Detroit, USA

Neuchâtel, le 11 avril 2019

Le Doyen, Prof. P. Felber

P. Fel

Abstract

This thesis is devoted to the study of the long-term behaviour of *Piecewise Deterministic Markov Processes* (PDMP). That is, a process $(X_t, I_t)_{t\geq 0}$, living in $\mathbb{R}^d \times E$ with E a finite space, where X is continuous and evolves between the jumps of I as

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

and I jumps according to

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

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

Here, for all $i \in E$, F^i denotes a vector field on $\mathbb{R}^d$ and for all $x \in \mathbb{R}^d$, $(a_{ij}(x))_{i,j\in E}$ is a rate matrix. In this thesis, we are mainly concerned with the following problem. Assume that for all $i \in E$,

$$F^i(0) = 0.$$

In other words, 0 is a common equilibrium point for the F^i. In particular, when the process X starts at 0, it remains there forever. A natural question is what is the behaviour of X when the starting point is not 0 but close to 0.

First, in a joint work with Michel Benaïm, we answer this question in a general context. We show, using stochastic persistence results, that the behaviour of (X, I) is mainly determined by the behaviour of the linearised process (Y, J), where $\dot{Y}_t = A^{J_t}Y_t$, A^i is the Jacobian matrix of F^i at 0 and J is the jump process with rates $(a_{ij}(0))$. We prove that under fairly general conditions, we can define a deterministic quantity Λ giving the exponential growth rate of Y as well as the behaviour of X near 0. More precisely, if $\Lambda < 0$, then $X_t \to 0$ exponentially fast with positive probability in a neighbourhood of 0, while if $\Lambda > 0$, the process (X, I) is *persistent* : it admits an invariant probability measure Π which does not confer mass to 0.

In a second joint work with Alexandru Hening, we show how we can apply this theory to answer a conjecture raised by Takeuchi et al. in 2006 on switching Lotka-Volterra prey-predator models. That is, we consider the case where there are two vector fields F^0, F^1 on $\mathbb{R}^2$ given by

$$F^i(x, y) = \begin{pmatrix} x(a_i - b_i y) \\ y(-c_i + d_i x) \end{pmatrix}.$$

We assume that the two vector fields have the same equilibrium point q in the positive quadrant and show that the average growth rate Λ at q is positive.

In a third work, I extend in a specific context the results obtained in the work with Michel Benaïm to the case where we can this time define two quantities $\Lambda^- < 0$ and

$\Lambda^+ > 0$, describing the exponential growth rates of Y. An application is given to Lorenz vector fields with switching.

The two remaining works of this thesis do not deal specifically with the above question, but are related to it. The first one is a joint work with Michel Benaïm and Tobias Hurth, where we give a slightly different condition from the ones existing for the exponential ergodicity of PDMP. It can in particular be applied in the case where the vector fields share a common equilibrium.

The final work is concerned with the discretisation in space of the PDMPs. That is, we look at a Markov chain on a finite grid, whose jump rates are related to the vector fields F^i. This Markov chain will hit 0 in finite time, and then remain in 0 forever. Thus, the Markov chain admits a quasi-stationary distribution. The purpose of the final chapter is to give some results on this quasi-stationary distribution when the size of the grid goes to infinity.

Keywords: Piecewise deterministic Markov processes; random switching; telegraph noise; Hörmander-bracket conditions; ergodicity; stochastic persistence ; Lyapunov exponents; epidemic models; SIS; population dynamics; Lotka-Volterra; Lorenz vector field; quasi-stationary distribution

Résumé

Cette thèse est dédiée à l'étude de *Processus de Markov Déterministes par Morceaux* (PDMP). C'est-à-dire, un processus (X, I) vivant sur $\mathbb{R}^d \times E$, où E est un ensemble fini, et X un processus continu, évoluant entre les sauts de I selon

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

Les sauts du processus I sont quant à eux gouvernés par le processus X :

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

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

Dans les équations ci-dessus, F^i désigne un champ de vecteurs de $\mathbb{R}^d$ et pour tout $x \in \mathbb{R}^d$, $(a_{ij}(x))_{i,j \in E}$ est une matrice de sauts. Dans cette thèse, nous nous intéressons essentiellement au problème suivant. Supposons que pour tout $i \in E$,

$$F^i(0) = 0.$$

Autrement dit, 0 est un point d'équilibre commun aux F^i, et si le processus X démarre à 0, il y restera pour toujours. Une question naturelle est donc de s'intéresser au comportement du processus X quand le point de départ n'est pas 0, mais en est proche.

Dans un premier travail publié conjointement avec Michel Benaïm, nous donnons une réponse à cette question dans un cadre général. Nous avons montré, en utilisant des résultats de persistance stochastique, que le comportement de X au voisinage de 0 se déduit essentiellement de celui du processus linéarisé (Y, J); où $\dot{Y}_t = A^{J_t} Y_t$, A^i est la matrice jacobienne de F^i en 0 et J est une chaîne de Markov sur E avec taux de sauts $(a_{ij}(0))$. Nous avons montré que nous pouvons définir une quantité réelle et déterministe Λ, qui donne le taux de croissance exponentiel de Y ainsi que le comportement de X près de 0. Plus précisément, si $\Lambda < 0$, alors $X_t \to 0$ exponentiellement vite avec probabilité positive si le point de départ est proche de 0, tandis que si $\Lambda > 0$, le processus est persistent : il admet une probabilité invariante qui ne donne pas de masse à 0.

Dans un deuxième travail, en collaboration avec Alexandru Hening, nous avons appliqué les résultats décrits ci-dessus pour répondre à une conjecture de Takeuchi et al. sur un système proie-prédateur de Lotka-Volterra en environnement fluctuant. Plus précisément, nous considérons le cas où il y a deux champs de vecteurs F^0 et F^2 sur $\mathbb{R}^2$ donnés par

$$F^i(x, y) = \begin{pmatrix} x(a_i - b_i y) \\ y(-c_i + d_i x) \end{pmatrix}.$$

Nous supposons que les deux champs de vecteur admettent le même point d'équilibre q dans le quadrant positif. Nous avons montré que dans ce cas, le taux de croissance Λ en q est positif. En particulier, le système ne peut pas converger vers q.

Dans un troisième travail, j'étends dans un certain contexte les résultats que nous avons obtenus avec Michel Benaïm au cas où nous pouvons cette fois définir deux taux de croissance avec des signes opposés. J'en donne une application à l'étude de champs de Lorenz modulés.

Les deux dernières parties de cette thèse ne traitent pas directement de la question ci-dessus, mais y sont reliées. L'une d'elle reprend un article que nous avons publié avec Michel Benaïm et Tobias Hurth, et dans lequel nous donnons une condition légèrement différente de celles existantes pour l'ergodicité des PDMP. Le résultat peut s'appliquer en particulier au cas où les champs de vecteurs ont un zéro commun.

Enfin, la toute dernière partie est dévolue à l'étude de la discrétisation en espace des PDMP considérés plus haut. Plus précisément, nous considérons une chaîne de Markov sur une grille de taille finie, dans les taux de sauts sont reliés au champs de vecteurs F^i. Cette chaîne de Markov touche 0 en temps fini, et ensuite n'en bouge plus. Ainsi, cette chaîne de Markov a une distribution quasi-stationnaire. Le but de cette dernière partie est l'étude du comportement asymptotique de cette distribution quasi-stationnaire lorsque la taille de la grille tend vers l'infini.

Mots clés: Processus de Markov Déterministes par Morceaux; modulation aléatoire; bruit télégraphe; condition de crochets à la Hörmander; ergodicité; persistence stochastique ; exposents de Lyapunov; modèles épidémiologiques; SIS; dynamique de population; Lotka-Volterra; champ de Lorenz; distribution quasi-stationnaire

Remerciements

La thèse est un travail personnel que l'on ne pourrait accomplir seul. Elle est faite de rencontres plus ou moins importantes, de nombreuses discussions, de collaborations, qui, au fil des années, forment un microcosme dont on se nourrit pour terminer son travail. C'est donc tout naturellement qu'avant d'entrer dans le vif du sujet, je souhaite remercier toutes les personnes qui m'ont aidé, de près ou de loin, à mener ma quête de théorèmes à bien.

Merci à mon directeur de thèse, Michel. Merci d'avoir soutenu le projet de cette thèse, merci de m'avoir donné ces questions passionnantes à résoudre. Merci de m'avoir toujours laissé une grande liberté, tant pour les thèmes abordés que pour les horaires de présence. Merci pour votre enthousiasme communicatif - "j'ai réfléchi hier soir en rentrant dans la voiture à notre petit problème et...", votre goût de la persévérance - " cent fois sur le métier remettre son ouvrage", votre définition des maths - "finalement, faire des maths, c'est mettre des trivialités bout à bout", vos anecdotes. Merci de m'avoir guidé - notamment pour la biblio - tout en me laissant très autonome. C'était vraiment un plaisir de travailler ces quatre années avec vous, de collaborer pour l'écriture de deux articles, et de toujours avoir des projets en cours. J'espère que nous pourrons les continuer, et que j'aurai l'occasion de revenir de temps à autre à Neuchâtel.

Merci Florent, ponte des PDMP, d'avoir accepté de rapporter et de participer au jury de ma thèse. Merci également pour l'invitation à la conférence sur les PDMP de Seillac en juin 2017, ainsi que pour nos discussions sur nos processus favoris lors de nos quelques rencontres en Suisse ou dans les pays de la Loire. Merci Denis et Nicolas, maîtres ès QSD, de participer au jury de ma thèse, et de me proposer de prolonger l'aventure à Nancy. J'ai hâte de relever de nouveaux défis avec vous ! Merci Felix de subir le rôle d'examinateur interne - aber kein Panik, es gibt kein Itô calculus in meiner Dissertation ! Thank you Sebastian for reporting on my thesis, and for your invitation to Davis and Snowbird in May - my wife hates you. Thank you also for all our interesting discussions on QSD, Lyapunov exponents and Roger Federer during your visits to Neuchâtel and my visit to the US. Finally, thank you Professor Yin for accepting to report on my thesis, it is a great honour for me.

I also wish to thank Alexandru for our collaboration and his proposition for a post-doctoral position in Boston. I hope we can still keep on working together on persistence issues.

Quiconque est déjà passé par l'Institut de Mathématiques de Neuchâtel peut en témoigner : l'ambiance qui y règne est excellente, et on ne peut le quitter qu'à regret - sauf le jeudi soir pour aller rejoindre sa chère et tendre. De plus, la tradition des remerciements dans la thèse s'y est développée de bien belle manière, et je vais tâcher d'y rendre hommage. Je vais commencer par mes collègues probabilistes, i.e, ceux qui font des maths intéressantes. Merci à mon grand frère de thèse, Carl-Erik, qui sévit maintenant dans un obscur métier en région parisienne. Merci pour tes conseils avisés, tant mathématiques qu'administratifs, pour nos longues discussions sur nos avenirs respectifs, le foot, et un peu le reste. Merci à ma grande soeur de thèse, Mireille. Merci pour ta bonne humeur permanente, ton rire déjà maintes fois encensé, tes conseils lors de mon arrivée ici. Dans la fratrie, je n'oublie bien sûr par Bastien, dont j'avais beaucoup apprécié la compagnie lors de mon tout premier stage neuchâtelois. Merci à Pierre d'avoir partagé mon bureau pendant un an, merci pour ton enthousiasme, pour tes jeux de mots. Thank you Clayton for bringing a little part of America in our office, it's refreshing ! Enfin, Hurth but not least, un grand merci à toi, Tobias, le plus grand joueur de hockey subaquatique que je connaisse. Merci pour toutes nos discussions sur les PDMP, pour notre collaboration sur l'article User-friendly condition, merci pour tes relectures incroyables de mes articles et de l'introduction de la présente thèse. C'est un très grand plaisir de partager un bureau avec toi, j'espère que nous pourrons continuer à travailler ensemble dans le futur.

Merci à Carsten et Lucas (prononcé Loukrass), les deux faux jumeaux du bureau d'en face - soit dit en passant, le seul attribué à des doctorants et qui donne sur le lac. Merci Lucas pour tes théories étonnantes et ta façon inimitable de parler français. Merci Carsten pour toutes nos discussions et nos parties de tennis - tu aurais au moins pu me laisser gagner une fois au nom de l'amitié franco-allemande. Merci pour tes gâteaux (et ceux de ta mère), ta bonne humeur immuable et de m'avoir mené en bateau. J'espère que tu survivras à ton exil en région zürichoise.

Merci à Luc et Aïsha (désolé Aïsha, j'ai mis dans cet ordre pour éviter le hiatus), les deux inséparables chamailleurs du bureau spectral. Par égards aux éventuelles âmes sensibles qui liraient ces lignes, nous ne détailleront pas ici toutes les tentatives de meurtre (au talon, au couteau, aux clés, au verre, à la poupée vaudou, au pain) que l'un a dû subir de l'autre. Je crois bien qu'il convient ici de parler de passion, dans toutes ses acceptions. Plus sérieusement, merci à Aïsha de représenter les minorités et d'être si sensible à la taquinerie. Merci à Luc, mon cher compatriote, pour ta verve, tes énigmes et nos nombreuses discussions. En parlant du bureau spectral, Hélène, tu pourras bientôt venir l'occuper sereinement sans risquer de subir un dommage collatéral.

Merci à Tom Empereur, dit Thomy. La mayonnaise a rapidement pris entre nous. Je sais que tu vas pleurer mon départ, mais tu as depuis rencontré ta grecquo-coréo-canadienne, Brigitte, avec qui je te souhaite une belle vie ! Merci à toi Tom de m'avoir accueilli quelques jours l'an dernier lors d'une période difficile pour moi. Je n'oublierai pas ta sollicitude à ces moments-là. Merci également pour tes autres invitations chez toi, pour tes bières belges, tes conseils pour Gand, Bruges et Bruxelles, et tout le reste.

Merci également aux deux autres algébristes belges glabres, Alain et Thiebout. Merci Alain pour la source intarissable d'anecdotes plus gouleyantes les unes que les autres dont tu nous gratifies. Merci Thiebout pour tes théories évoluées et tes spaghettis bolognaises.

Merci Valentin, alias Jésus, pour tes punchlines comme "le gras, c'est la vie", "à la bouffe", "vous êtes tous gros", "salut les moches", tes horaires précis, et tes factorisations.

Merci Alexandre "zumzum19" d'avoir considérablement retardé la rédaction de cette thèse grâce à tes venues incessantes dans mon bureau ponctuées de "j'ai une petite question de théorie de la mesure..." Merci pour ta bonne humeur, tes attitudes de beau gosse et tes "Aaaadios". Je ne te remercie pas de m'enfumer quand je t'accompagne pendant que tu vas cloper.

Merci Laura d'avoir gagné presque tous les chocolats du calendrier de l'Avent, ça m'a empêché de grossir.

Merci Radu de m'avoir fait découvert la Porn Star.

Merci également aux cryptographes Alberto, Elisa, Giulia, Christina, Relinde et Alessio.

Merci Katie d'apporter de la douceur et de la gentillesse dans ce monde de brutes.

Merci Ana pour ton humour et tes conseils sur l'Angleterre. Je te souhaite une bonne suite à Cambridge.

Merci Laurent de m'avoir démontré qu'on pouvait garder son bonnet à l'intérieur et de m'accompagner parmi les gens qui apportent leur bouffe de la maison. Merci également pour ton aide pour les animations pour la présentation publique.

Merci Hugues pour ton enthousiasme, c'était vraiment sympa d'être ton assistant. Et merci d'avoir laissé trainer ta thermos sur mon bureau pendant 18 mois.

Merci à Sasha pour les superbes apéros suivants les colloques du mardi.

Merci Rémi de nous laisser des places au saloon pour les pauses de midi.

Merci Bruno pour ton aide lors de l'assistanat de CDI 1.

Merci également aux plus anciens que je n'ai pas encore cités : merci Régis pour nos quelques mois de colocation à Arcueil, ce sont vraiment de bons souvenirs - ah, ce magret de canard - haricots verts - pommes de terre grillées directement cuisiné après le marché ! Merci à David pour ton humour si particulier et nos discussions très intéressantes, ainsi que pour nos quelques parties de tennis. Merci également à Pierre-Nicolas, Basile, Yoann.

Merci à Christine pour son efficacité et pour nous guider dans la jungle de l'administration.

Neuchâtel ne se résume pas à l'Institut de Mathématiques, et je voudrais remercier du fond du coeur la famille Babaiantz pour leur accueil et leur générosité durant ses quatre années de thèse (et même plus). Mes pensées vont en particulier à Nadine, qui n'est malheureusement plus là pour assister à la fin de cette aventure. Merci pour ton accueil il y a six ans déjà, merci pour ton enthousiasme, ta force, ton sourire, ta cuisine, la main que tu avais sur le coeur - tout cela me manque terriblement. Merci Olivier pour ton affabilité, ta présence, nos discussions. Merci aussi à Raoul, Iris, Aymone, Arthur, Cyril (n'oublie pas les cocobolides) et la dernière venue, Ada. Un immense merci à vous tous.

Une partie non négligeable de cette thèse ayant été rédigée dans le train, je remercie la SNCF et les CFF d'avoir mis à ma disposition des trains toutes les trente minutes pour faire mes allers-retours entre Neuchâtel et Sélestat.

Merci à mes amis, les anciens et fidèles Guillaume et Julien, ceux du Lycée, Antoine et Jérémie, ceux de prépa, Lucas, Matthieu, Pierre-Antoine et Simon, ceux de l'ENS, Simon, Ludovic, Jessica et Romain, Claire et Alain, Benjamin, Lilian. En parlant de l'ENS, merci à Alain Trouvé, c'est lui qui m'orienta vers Michel pour un stage de master en 2013 - finalement, le premier germe de cette thèse.

Merci à ma famille, mes parents qui m'ont toujours soutenu dans les études, notamment lors des nombreux déménagements qu'elles ont engendrés (le piano vous salue), ma fratrie, Benjamin, j'ai toujours pu profiter de tes conseils avisés, et Cornélia, pour m'avoir appris à compter les éléphants sur les albums de Babar, et les pièces rapportées, Laurianne

et Sylvain, merci également à ma nièce Anaée de me donner des petits calculs récréatifs pendant les vacances.

Merci également à la famille Mosser, le pendant alsacien de la famille Babaiantz ! Merci pour votre accueil pendant les trois années précédant le début de cette thèse. Merci Marie-Odile pour ta grande générosité, pour ta vivacité, pour ta bonne humeur, merci Gilbert pour les même choses, merci la Grosse de m'avoir donné un aperçu de ce que c'est une petite soeur.

Enfin, merci à toi, Chloé. Merci de partager ta vie avec moi, merci d'avoir supporté mes trop nombreuses absences et mes blagues de maths, merci d'être toujours là pour moi, merci d'accepter la vie de romanichels qui nous attend pour les années à venir, merci pour ta bonne humeur et ton sourire, je ne pourrai pas vivre sans.

Cette thèse a été partiellement financée par les bourses du Fonds National Suisse pour la Recherche 200021_175728 et 200021_163072

Notations

The following notations will be used throughout the thesis. Let (M, d) be a metric space, $x \in M$, $t \geq 0$, n, m integers, and $X = (X_t)_{t\geq 0}$ a Markov process.

- $\mathcal{B}(M)$: Borel sigma field of M
- For $A \in \mathcal{B}(M)$, $\mathbb{1}_A$ is the indicator function of A
- $L_b(M)$: space of measurable bounded functions $f : M \to \mathbb{R}$
- $C^n(M)$: space of functions $f : M \to \mathbb{R}$ of class C^n
- $C^n_c(M)$: space of functions of class C^n with compact support
- $C_b(M)$: space of bounded continuous maps $f : M \to \mathbb{R}$
- $\mathcal{P}(M)$: set of probability measures on $(M, \mathcal{B}(M))$
- δ_x : Dirac mass at x
- For $f \in L_b(M)$ and $\mu \in \mathcal{P}(E)$, $\mu f = \int_M f d\mu$.
- $\mathbb{P}_x$: Law of X starting at point x
- $\mathbb{E}_x$: expectation associated to $\mathbb{P}_x$
- For $f \in L_b(M)$ and $A \in \mathcal{B}(M)$, $P_t f(x) = \mathbb{E}_x(f(X_t))$, $P_t(x, A) = \mathbb{P}_x(X_t \in A)$
- $\mathcal{P}_{inv}$: set of invariant probability measures of X
- $\mathcal{P}_{erg}$: set of ergodic probability measures of X
- If $M' \in \mathcal{B}(M)$, $\mathcal{P}_{inv}(M')$: set of invariant probability measures of X such that $\mu(M') = 1$
- $\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \ : x_i \geq 0, \forall i = 1, \ldots, n\}$
- $\mathbb{R}^n_{++} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \ : x_i > 0, \forall i = 1, \ldots, n\}$
- $M_n(\mathbb{R})$: set of real matrices of size $n \times n$
- $GL_n(\mathbb{R})$: set of invertible real matrices of size $n \times n$
- For $F : \mathbb{R}^n \to \mathbb{R}^m$, $DF(x)$: differential of F at point x
- a.s : almost surely
- càdlàg : right continuous with left limit

Contents

Introduction

To die or not to die? This is the central question that the processes studied in this thesis ask themselves. For example, will a population living in a random environment survive or eventually become extinct ? Will a disease evolving randomly persist or disappear ?

As a motivating example, consider a disease spreading in a population which is divided in several groups such that the infectiveness and the recovery rate of an individual depends on which group they belong to. For example, we can imagine that old individuals are more likely to be infected, or that the disease affects more males than females : this will give rise to different age or gender categories. For a more detailed presentation, one could read the introduction of [ARBMW14]. Let $d \geq 1$ be the number of groups, and for each $i = 1, \ldots, d$, let $x_i(t) \in [0,1]$ denote the proportion of infected individuals in group i at time $t \geq 0$. We can model the evolution of the disease by a differential equation:

$$\frac{dx(t)}{dt} = F(x(t)), \tag{1}$$

where $x = (x_1, \ldots, x_d)$ and $F : [0,1]^d \to \mathbb{R}^d$ is a globally integrable vector field. For example, the famous model of Lajmanovich and Yorke [LY76] for gonorrhea is given by :

$$\frac{dx_i}{dt} = (1 - x_i)(\sum_{j=1}^{d} C_{ij} x_j) - D_i x_i \, , \; i = 1, \ldots d, \tag{2}$$

where C_{ij} and D_i are positive constants, representing respectively the transmission rate from an individual of group j to an individual of group i and the recovery rate of an individual of group i. We assume that once the disease has disappeared - that is, $x_i = 0$ for all $i = 1, \ldots, d$ - it cannot reinfect the population. Mathematically, this assumption is translated into $F(0) = 0$. In other words, 0 is an *equilibrium* of the differential equation (1). In this context, it is called the *Disease Free Equilibrium.* Under fairly general conditions on F, it can be proven that exactly one of the following holds : either this equilibrium is globally asymptotically stable, meaning that every trajectory converges to 0, or there exists another equilibrium $x^* \in [0,1]^d$ with positive coordinates whose basin of attraction is $[0,1]^d \setminus \{0\}$ (see e.g. [LY76], Hirsch [Hir94], and Theorem 3.8 in Chapter 3). When x^* exists, it is called an *Endemic Equilibrium*, and it is reflecting the fact the the disease is eventually becoming ingrained in the population at some proportion.

Now, to take into account the haphazardness of the environment, we will assume that the vector field in (1) is sometimes randomly switched with another vector field. That is, we consider a family of vector fields $(F^i)_{1 \leq i \leq m}$ sharing the properties of F and we consider the differential equation

$$\frac{dX(t)}{dt} = F^{I_t}(X(t)), \tag{3}$$

where $(I_t)_{t \geq 0}$ is a *random process* with càdlàg paths in $\{1, \ldots, m\}$. More precisely, X is a continuous process satisfying (3) between the jumps of I. We only consider the case where

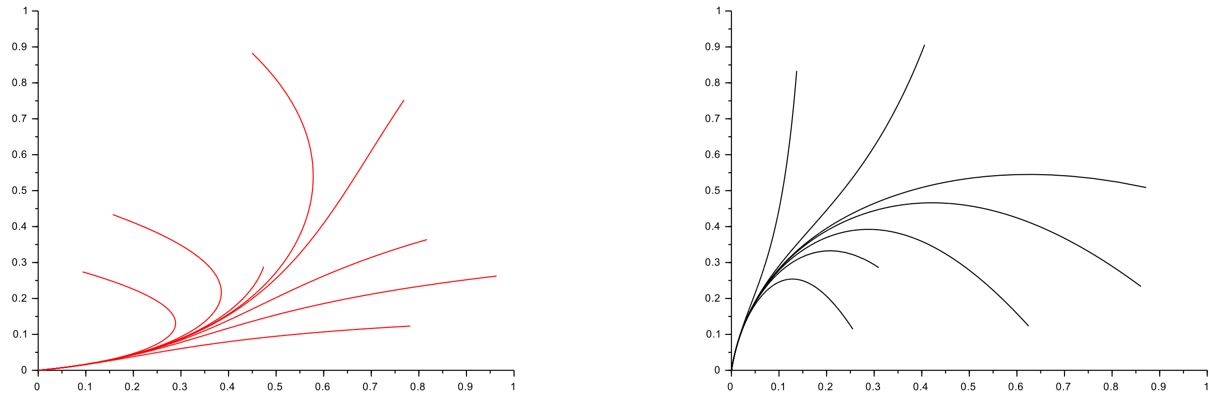


Figure 1: Phase portrait of F^0 and F^1

$(I_t)_{t\geq 0}$ is a continuous time Markov chain on $\{1,\dots,m\}$. A natural question is whether one can still say something about the long-term behaviour of X. One of the main results of this thesis is to give a positive answer to that question (see Theorem 1.28 in Section 1.6), and can be formulated as follows :

Theorem 0.1. *There exists a real number* Λ *such that*

1. *If* $\Lambda < 0$*, then* $\lim_{t\to\infty} X_t = 0$ *almost surely for all initial conditions,*

2. *If* $\Lambda > 0$*, there exists a unique invariant probability measure* Π *for the process* (X, I) *such that* $\Pi(\{0\} \times \{1,\dots,m\}) = 0$*. Furthermore, the law of* (X_t, I_t) *converges to* Π *provided* $X_0 \neq 0$*.*

This theorem belongs to the class of *stochastic persistence* results. It is symptomatic of the kind of results that are obtained in this thesis, and that are summed up in Section 1.6 below. The following example, taken from Chapter 3 illustrates the fact that Λ can be positive *even though the Disease Free Equilibrium is globally asymptotically stable for all the* F^i.

Example 0.1 (Fluctuations may promote infection)**.** We consider two vector fields F^0 and F^1 as above, with

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

and

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

It is easily seen that the eigenvalues of the Jacobian matrices of F^0 and F^1 at zero are negative, thus 0 is globally asymptotically stable for both F^0 and F^1 : in each environment taken individually, the disease disappears. Phase portraits of F^0 and F^1 are given in Figure 1. Now let $(I_t)_{t\geq 0}$ be a Markov chain on $\{0,1\}$ jumping from 0 to 1 and from 1 to 0 at the same rate $\beta > 0$. We show in Chapter 3 that when β is sufficiently large, Λ is positive : when the environment quickly switches between the states 0 and 1, the disease persists in the population. This is illustrated in Figure 2.

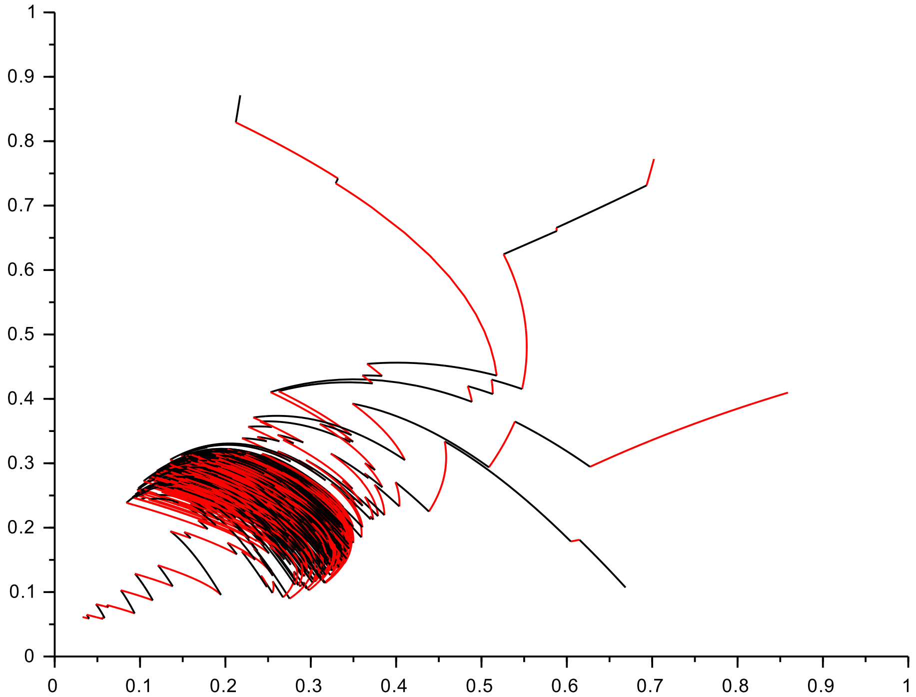


Figure 2: Some trajectories of X_t for $\beta = 20$

The process (X, I) described above is a particular case of a *Piecewise Deterministic Markov Process* (PDMP for short). That is, a Markov process evolving deterministically between some random jumps occurring at random times. The most important part of this thesis deals with the persistence of such processes: if a PDMP is modelling a population, what can be said about the long-term behaviour of the system ? As illustrated by the above example, the behaviour of the PDMP can be fundamentally different from the behaviours obtained under each fixed environment. Several other examples will be given in this thesis.

Chapter 1

Fundamental tools and description of the main results

This chapter gives a presentation of the processes studied in this thesis, and sums up the main original results obtained. It is organised as follows. First, in Section 1.1, we recall briefly some basic tools and notations used in the theory of Markov processes. In Section 1.2, we give the general definition of a PDMP and some examples. Subsection 1.2.4 is devoted to the more particular case of PDMP generated by random switching of vector fields. Then we introduce in Section 1.3 the theory of stochastic persistence developed by Michel Benaïm and show how it applies to a particular PDMP. Section 1.4 is devoted to a brief introduction to Random Dynamical Systems and Lyapunov exponents. In Section 1.5, we recall the concept of quasi-stationary distribution and recall some of the recent results of Nicolas Champagnat and Denis Villemonais. Finally, in Section 1.6, we describe more precisely the main results obtained in this thesis.

Contents

1.1 Definitions and tools for Markov Processes

In this first section, we give some notation, concepts and results from the theory of Markov processes that will be used throughout this thesis. We shall assume from this point that the reader is aware of this theory, and the goal is neither to give an introduction to Markov processes nor to be exhaustive. The main objective is to clear up ambiguities that may arise from non-unique terminology (generator, Feller, etc...)

We consider a continuous-time Markov process $(X_t)_{t\geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ and with càdlàg paths in a locally compact metric space (M,d). For ν a probability measure on M, we set, as usual, $\mathbb{P}_\nu$ for the law of the process X on the space of càdlàg functions $\mathbb{D}(\mathbb{R}_+, M)$ with initial distribution ν and $\mathbb{E}_\nu$ for the associated expectation. If $\nu = \delta_x$ for some $x \in M$, we write $\mathbb{P}_x$ for $\mathbb{P}_{\delta_x}$.

1.1.1 Semigroup

We start with a definition. We denote by $L_b(M)$ the set of bounded measurable functions $f : M \to \mathbb{R}$, and for $f \in L_b(M)$, we set

$$\|f\| = \|f\|_\infty = \sup_{x\in M} |f(x)|.$$

Definition 1.1. *For every $t \geq 0$, we consider the operator $P_t : L_b(M) \to L_b(M)$ such that, for all function $f \in L_b(M)$ and all $x \in M$,*

$$P_t f(x) = \mathbb{E}_x\left(f(X_t)\right).$$

Sometimes one also see P_t is a Markov kernel : for all $x \in M$, and all $A \in \mathcal{B}(M)$, we write

$$P_t(x, A) = P_t \mathbb{1}_A(x) = \mathbb{P}_x(X_t \in A).$$

The Markov property of X implies the so called *Chapman-Kolmogorov equations*

$$P_{t+s} = P_t \circ P_s = P_s \circ P_t,$$

$$P_0 = Id.$$

The family $(P_t)_{t\geq 0}$ is thus a semigroup of operators, called the Markov semigroup associated to $(X_t)_{t\geq 0}$. For a distribution ν on M, we let νP_t be the measure defined by

$$\nu P_t f = \int_M P_t f(x) d\nu(x),$$

for all $f \in L_b(M)$. Thus νP_t is the law of the Markov process X at time t, knowing that X_0 is distributed according to ν.

Definition 1.2. *A probability measure μ on M is an* invariant probability measure *or* invariant distribution *if*

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

1.1.2 Generator

The *infinitesimal generator* of the semigroup $(P_t)_{t\geq 0}$ is an operator L whose domain is given by

$$\mathcal{D}(L) = \left\{ f \in L_b(M) : \exists g \in L_b(M), \lim_{t\to 0} \left\| \frac{P_t f - f}{t} - g \right\| = 0 \right\}.$$

For $f \in \mathcal{D}(L)$, we define $Lf(x)$ for all $x \in M$ as

$$Lf(x) = \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t}.$$

It describes the infinitesimal evolution of the process: since for all $t, h \geq 0$,

$$\mathbb{E}_x \left[f(X_{t+h}) - f(X_t) | \mathcal{F}_t \right] = P_h f(X_t) - f(X_t), \quad \mathbb{P}_x - \text{a.s.}$$

one can write

$$\mathbb{E}_x \left[f(X_{t+h}) - f(X_t) | \mathcal{F}_t \right] = hLf(X_t) + o(h).$$

From the definition, $Lf(x)$ can be seen as the derivative at time 0 of $P_t f(x)$. Actually, it describes the derivatives of $P_t f(x)$ at every point t, as stated in the next proposition (see e.g. [Dav93, Prop 14.10]).

Proposition 1.1. *Let $f \in \mathcal{D}(L)$. Then*

1. *For all $t \geq 0$, $P_t f \in \mathcal{D}(L)$ and*

$$\frac{dP_t f(x)}{dt} = LP_t f(x) = P_t Lf(x). \tag{1.1}$$

2. *For all $t \geq 0$,*

$$P_t f(x) = f(x) + \int_0^t P_s Lf(x) ds. \tag{1.2}$$

Equation (1.2) can be rewritten as

$$\mathbb{E}_x(f(X_t)) = f(x) + \mathbb{E}_x \left[\int_0^t Lf(X_s) ds \right],$$

and is known as *Dynkin's formula*. It turns out that thanks to the Markov property, the following stronger proposition is true (see e.g. [Dav93, Prop 14.13]).

Proposition 1.2. *For all $f \in \mathcal{D}(L)$, for all $x \in M$, the process*

$$M_t^f(x) = f(X_t) - f(x) - \int_0^t Lf(X_s) ds$$

is a $(\mathcal{F}_t)_{t\geq 0}$ - martingale under $\mathbb{P}_x$.

This leads to an extension of the definition of the generator :

Definition 1.3. *The* domain of the extended generator $\mathcal{D}(\bar{L})$ *is the set of functions $f \in L_b(M)$ such that there exists $g \in L_b(M)$ for which the process*

$$M_t^{f,g}(x) = f(X_t) - f(x) - \int_0^t g(X_s) ds \tag{1.3}$$

is an $(\mathcal{F}_t)_{t\geq 0}$ - local martingale under $\mathbb{P}_x$

It can be proven (see e.g. [Dav93, Definition 14.15]) that if for two functions g_1 and g_2, $M^{f,g_1}(x)$ and $M^{f,g_2}(x)$ are local martingales under $\mathbb{P}_x$, then $g_1 = g_2$ except on a set A of *zero potential*, that is a set such that

$$\mathbb{P}_x\left[\int_0^{+\infty} \mathbb{1}_A(X_s)ds = 0\right] = 1, \quad \forall x \in M$$

In other words, the process X is 'never' in A. This legitimates the notation

$$\bar{L}f = g,$$

which completes the definition of the extended generator. When we want to emphasise the difference between the infinitesimal generator and the extended one, we sometimes refer to L as the *strong generator*. By the previous proposition and the definition of the extended generator, the following lemma is obvious :

Lemma 1.1. *We have $\mathcal{D}(L) \subset \mathcal{D}(\bar{L})$, and for all $f \in \mathcal{D}(L)$, $Lf = \bar{L}f$.*

The infinitesimal generator can be used to find an invariant probability measure, thanks to the next proposition. We say that a class of functions $\mathcal{C} \subset L_b(M)$ is *separating* if for every $\mu, \nu \in \mathcal{P}(M)$,

$$\mu f = \nu f, \quad \forall f \in \mathcal{C} \Rightarrow \mu = \nu.$$

We introduce the set $\mathbb{B}_0$ as

$$\mathbb{B}_0 = \{f \in L_b(M) \ : \lim_{t\to 0} \|P_t f - f\| = 0\}.$$

Proposition 1.3. *([EK86, Proposition 9.2 p. 239]) Assume that $\mathbb{B}_0$ is separating and let μ be a probability measure on M. Then μ is invariant if and only if, for all $f \in \mathcal{D}(L)$, $\mu Lf = 0$.*

In general, it is hard, if not intractable to find the domain $\mathcal{D}(L)$. However, it is often possible to exhibit a sufficiently large subset of the domain on which one can try the criterion of the above proposition. This is the sense of the following definition.

Definition 1.4. *A subset D of $\mathcal{D}(L)$ is a* core *for $(\mathcal{D}(L), L)$ if for all $f \in \mathcal{D}(L)$, there exists a sequence $(f_n)_{n\geq 0}$ of functions in D such that* $\lim_{n\to\infty} f_n = f$ *and* $\lim_{n\to\infty} Lf_n = Lf$.

With this definition, we have the following weaker statement.

Proposition 1.4. *([EK86, Proposition 9.2 p. 239]) Assume that $\mathbb{B}_0$ is separating, D is a core for $(\mathcal{D}(L), L)$ and let μ be a probability measure on M. Then μ is invariant if and only if, for all $f \in D$, $\mu Lf = 0$.*

1.1.3 Feller property

If there is one notion which carries its share of ambiguities, it is that of *Feller property*. It changes from author to author, and it has numerous sisters : weak Feller, strong Feller, asymptotically strong Feller, and so on. Here we give the most widely spread definition of the Feller property and the one used in this thesis, which is weaker ! Recall that C_0 is the set of continuous functions $f : M \to \mathbb{R}$ vanishing at infinity : for all $\varepsilon > 0$, there exits $K \subset M$ compact such that $|f(x)| < \varepsilon$ for all $x \notin K$.

Definition 1.5. *We say that a Markov semigroup* $(P_t)_{t\geq 0}$ *is* Feller *or has the* Feller property *if :*

1. *For all* $t \geq 0$, $P_t C_0 \subset C_0$*;*

2. *For all* $f \in C_0$, $\lim_{t\to 0} \|P_t f - f\| = 0$.

We say that a Markov process is Feller if its associated semigroup is Feller. When we want to emphasise that the semigroup preserves C_0, we shall say that it is C_0 - Feller. Since C_0 is separating, the above definition and Proposition 1.3 immediately yield :

Proposition 1.5. *Let* $(P_t)_{t\geq 0}$ *be a Feller semigroup with strong generator* $(\mathcal{D}(L), L)$ *and* $\mu \in \mathcal{P}(M)$*. Then* μ *is invariant if and only if, for all* $f \in \mathcal{D}(L)$, $\mu L f = 0$.

We now give the weaker notion of Feller used in this thesis. We denote by $C_b(M)$ the set of bounded continuous functions $f : M \to \mathbb{R}$.

Definition 1.6. *We say that a Markov semigroup* $(P_t)_{t\geq 0}$ *is* C_b - Feller *if, for all functions* $f \in C_b(M)$*, the map* $(x,t) \mapsto P_t f(x)$ *is continuous.*

We prove that this definition is indeed weaker than Definition 1.5.

Lemma 1.2. *Every* C_0 *- Feller semigroup is* C_b *- Feller.*

Proof We slightly adapt the proof given in [Sch98, Theorem 3.2] Assume that $(P_t)_{t\geq 0}$ is a C_0 - Feller semigroup and let $f \in C_0$. Then, the map $(x,t) \to P_t f(x)$ is continuous. Indeed, for all $0 \leq s \leq t$ and all $x, y \in M$, the semigroup property implies that

$$\begin{aligned} |P_{t+s} f(x) - P_t f(y)| &\leq |P_{t+s} f(x) - P_t f(x)| + |P_t f(x) - P_t f(y)| \\ &\leq \|P_s f - f\| + |P_t f(x) - P_t f(y)|, \end{aligned}$$

and

$$\begin{aligned} |P_{t-s} f(x) - P_t f(y)| &= |P_{t-s+s} f(y) - P_{t-s} f(x)| \\ &\leq \|P_s f - f\| + |P_{t-s} f(x) - P_{t-s} f(y)|, \end{aligned}$$

which gives the result from the two points of Definition 1.5. Now we prove as in [Sch98, Theorem 3.2] that if $f \in C_b$, there exists a sequence of maps $f_n \in C_0$ such that $(x,t) \mapsto P_t f_n(x)$ converges uniformly on compact set towards $(x,t) \mapsto P_t f(x)$. Let $\varphi_n \in C_0$ such that $0 \leq \varphi_n \leq 1$ and the sequence φ_n increases monotonically to 1 and set $f_n = \varphi_n f$. Then, for all (x,t),

$$|P_t f(x) - P_t f_n(x)| \leq \|f\| P_t(1 - \varphi_n)(x).$$

Now by dominated convergence, for all (x,t), $\lim_{n\to\infty} P_t(\varphi_n)(x) = 1$. By Dini's Theorem, the convergence is uniform over compact sets of $M \times \mathbb{R}_+$, thus $(x,t) \mapsto P_t f_n(x)$ converges uniformly on compact sets towards $(x,t) \mapsto P_t f(x)$. Since $f_n \in C_0$ for every n, it implies by the beginning of the proof that $(x,t) \mapsto P_t f(x)$ is continuous. □

One of the important features of the Feller property is to guarantee the existence of an invariant probability measure provided the state space is compact:

Proposition 1.6. *Let* $(P_t)_{t\geq 0}$ *be a* C_b *- Feller semigroup on some compact metric space* M*. Then* $(P_t)_{t\geq 0}$ *admits at least one invariant probability measure.*

Proof We give the proof using the so-called Krylov - Bogoliubov procedure. Fix some $x \in M$, and for all $t > 0$, define the mean empirical measure as

$$\mu_t = \frac{1}{t}\int_0^t P_s(x,\cdot)ds.$$

That is, for every bounded measurable function f,

$$\mu_t f = \frac{1}{t}\int_0^t P_s f(x)ds.$$

As M is compact, the sequence of measures $(\mu_t)_{t>0}$ is tight. We show that every weak-limit point of this sequence is an invariant probability measure of $(P_t)_{t\geq 0}$. Let μ be such a limit point and $(t_n)_{n\geq 0}$ an increasing sequence going to infinity such that $\mu_{t_n} \Rightarrow \mu$. For $u > 0$ and $f \in C_b(M)$, the semigroup property implies that

$$\mu_{t_n} P_u f = \mu_{t_n} f - \frac{1}{t_n}\int_0^u P_s f(x)ds + \frac{1}{t_n}\int_{t_n}^{u+t_n} P_s f(x)ds.$$

Since for all $s \geq 0$, $\|P_s f(x)\| \leq \|f\|$, the right - hand side of the above equality goes to μf as n goes to infinity. On the other hand, since the semigroup is C_b - Feller, $\lim \mu_{t_n} P_u f = \mu P_u f$. Thus $\mu P_u f = \mu f$ for all continuous bounded functions f and all $u \geq 0$, meaning that μ is an invariant probability measure of $(P_t)_{t\geq 0}$. □

Remark 1.1. *In the above proof, compactness of M is only used to ensure the tightness of the sequence of the mean empirical measures. Therefore, if one can prove this tightness, the conclusion still holds. This is for example the case when there exists a* Lyapunov function, *i.e a positive function V going to infinity at infinity and such that, formally,*

$$LV \leq -\lambda V + C,$$

for some nonnegative constants λ, C (see the next section and Section 1.3).

For the remainder of this thesis, if not otherwise specified, Feller will mean C_b - Feller.

1.1.4 Doeblin condition and ergodicity

In the previous section, we saw that the Feller property and the compactness of the state space yield the existence of an invariant probability measure. It is natural then to wonder if this invariant probability measure is unique, and if the process converges in some sense to it. One of the most useful tools to prove uniqueness and convergence is the *Doeblin condition*, combined with the existence of a Lyapunov function when the state space is noncompact. This section is devoted to the introduction of the notions of Doeblin condition, Doeblin set, Doeblin point and small set, and the statement of the principal theorems involving these concepts.

Detour through the discrete chain

In this section, we recall some well-known facts on discrete time Markov chains. Let $\mathcal{P}$ be a Markov kernel on some metric space E endowed with its Borel sigma field $\mathcal{B}(E)$. We let $\mathcal{P}^n$ denote the associated n - step transition kernel . We have the following classical definitions :

Definition 1.7. *A set $C \subset E$ is called a* petite set *if there exists a non-trivial measure ν, and a sequence of positive numbers $(q_n)_{n\geq 0}$ with $\sum q_n = 1$ such that, for every $x \in C$, for every $A \in \mathcal{B}(E)$,*

$$\sum_{n\geq 0} q_n \mathcal{P}^n(x, A) \geq \nu(A).$$

Definition 1.8. *A set $C \subset E$ is called a* small set *if there exists a non-trivial probability measure ν, an integer $m \geq 1$, and $0 < \varepsilon < 1$ such that, for every $x \in C$, for every $A \in \mathcal{B}(E)$,*

$$\mathcal{P}^m(x, A) \geq \varepsilon\nu(A).$$

Whenever C is nonempty, we shall say that $\mathcal{P}$ satisfies a Doeblin condition.

Let $\mu, \nu \in \mathcal{P}(E)$. The total variation distance between μ and ν is defined as

$$d_{TV}(\mu, \nu) = \frac{1}{2}\|\mu - \nu\|_{TV} = \sup_{A\in\mathcal{B}(E)} |\mu(A) - \nu(A)|.$$

It can also be expressed via functions :

$$\|\mu - \nu\|_{TV} = \sup\{\mu f - \nu f : \; f \in L_b(E), \|f\| \leq 1\},$$

and via its dual formulation :

$$d_{TV}(\mu, \nu) = \inf_{X\sim\mu, Y\sim\nu} \mathbb{P}(X \neq Y).$$

Definition 1.9. *A set C is a* Doeblin set *if there exists an integer $m \geq 1$, and $0 < \varepsilon < 1$ such that, for every $x, x' \in C$,*

$$d_{TV}(\mathcal{P}^m(x, \cdot), \mathcal{P}^m(x', \cdot)) \leq 1 - \varepsilon.$$

The above definitions are linked :

Lemma 1.3. *Every small set is Doeblin and petite. If the chain is irreducible and aperiodic, then every Doeblin set and every petite set is small.*

The fact that a small set is petite is obvious from the definitions. For proofs of the other points, the reader is referred to [DMPS18]. The interest of these notions is that they lead to an anthology of results on ergodicity of the chain. We start with the following, which is part of Proposition 6.1.9 in [Duf97].

Proposition 1.7. *If $\mathcal{P}$ admits a petite set, then $\mathcal{P}$ admits at most one invariant probability measure.*

The next theorem is among of the most famous ones on Markov chains. A proof can be found for example in [DMPS18].

Theorem 1.1. *Assume that E is a Doeblin set. Then $\mathcal{P}$ admits a unique invariant probability measure π, and there exists $\rho \in (0, 1)$ such that, for all $n \geq 0$ and all $\xi \in \mathcal{P}(E)$,*

$$\|\xi\mathcal{P}^n - \pi\|_{TV} \leq \|\xi - \pi\|_{TV}\rho^n.$$

If the conclusion of the above theorem holds, we shall say that $\mathcal{P}$ is *uniformly geometrically ergodic*. When the state space is noncompact, it is in general hard if not impossible to prove that E is a Doeblin set. However, if compact subsets of E are small sets, and if we control the behaviour of $\mathcal{P}$ at the boundary of E with a Lyapunov function, then we still get geometric ergodicity of the process, but it would no longer be uniform. This is the content of the famous Harris Theorem that we state now. We reproduce the statement of Hairer and Mattingly and refer to [HM11] for a proof. For a function $V : E \to [0, \infty)$, we introduce the weighted norm on $L_b(E)$ defined for all f by

$$\|f\|_V = \sup_{x \in E} \frac{|f(x)|}{1 + V(x)},$$

and the associated weighted total variation norm :

$$\|\mu - \nu\|_V = \sup\{\mu f - \nu f : \ f \in L_b(E), \|f\|_V \leq 1\}.$$

Theorem 1.2. *Assume the following :*

1. *There exist a function $V : E \to [0, \infty)$, and constants $\gamma \in (0, 1)$, $K \geq 0$ such that, for all $x \in E$,*
$$\mathcal{P}V(x) \leq \gamma V(x) + K.$$
2. *There exists $R \geq \frac{2K}{1-\gamma}$ such that the set $A = \{x \in E : \ V(x) \leq R\}$ is a small set.*

Then, $\mathcal{P}$ admits a unique invariant probability measure π. Moreover, πV is finite and there exist constants $\rho \in (0, 1)$ and $C \geq 0$ such that, for all $n \geq 0$, and every $f \in L_b(E)$ with $\|f\|_V \leq 1$,

$$\|\mathcal{P}^n f - \pi f\|_V \leq C\rho^n \|f - \pi f\|_V.$$

In particular, for all $x \in E$;

$$\|\delta_x \mathcal{P}^n - \pi\|_V \leq C' \rho^n (1 + V(x) + \pi V).$$

Remark 1.2. *A function V satisfying point* 1. *in the above theorem is called a* Lyapunov function *(for $\mathcal{P}$).*

Proof We only prove the the last inequality, the beginning of the theorem is [HM11, Theorem 1.2]. Let $x \in E$ and $f \in L_b(E)$ with $\|f\|_V \leq 1$. Then

$$\begin{aligned} |\mathcal{P}^n f(x) - \pi f| &\leq \|\mathcal{P}^n f - \pi f\|_V (1 + V(x)) \\ &\leq C\rho^n \|f - \pi f\|_V (1 + V(x)) \\ &\leq C\rho^n (\|f\|_V + \pi f)(1 + V(x)). \end{aligned}$$

Now

$$\mathcal{P}f \leq \|f\|_V (1 + PV) \leq 1 + K + \gamma V,$$

which implies, by invariance of π, $\pi f \leq 1 + K + \pi V$, hence the result. $\square$

Remark 1.3. *The second part of the theorem can also be proven using V - Dobrushin coefficient, see [DMPS18].*

Return to the continuous time

We now go back to a continuous-time Markov semigroup $(P_t)_{t\geq 0}$ on a metric space (M, d). We start by giving the counterpart of the above definitions in this setting.

Definition 1.10. *A set $C \subset M$ is called a* petite set *if there exists a non-trivial measure ν and a probability measure γ on $\mathbb{R}_+$ such that, for every $x \in C$ and every $A \in \mathcal{B}(M)$,*

$$\int_0^{+\infty} P_t(x, A) d\gamma(t) \geq \nu(A).$$

We derive immediately from the discrete case the following lemma :

Lemma 1.4. *Assume $(P_t)_{t\geq 0}$ admits a petite set C. Then there exists at most one invariant probability measure for $(P_t)_{t\geq 0}$.*

Proof Let R be the Markov kernel defined by $R(x, \cdot) = \int_0^{+\infty} P_t(x, \cdot) d\gamma(t)$. Then C is a petite set for R, thus R admits at most one invariant probability measure. Now it is easily seen that every invariant probability measure of $(P_t)_{t\geq 0}$ is also invariant for R, hence the result. □

Definition 1.11. *A set $C \subset M$ is called a* small set *if there exists a non-trivial probability measure ν, a positive t, and $0 < \varepsilon < 1$ such that, for every $x \in C$, for every $A \in \mathcal{B}(M)$,*

$$P_t(x, A) \geq \varepsilon \nu(A).$$

Definition 1.12. *A set $C \subset M$ is a* Doeblin set *if it is a Doeblin set for the Markov kernel P_t for some $t > 0$.*

As in the discrete case, it is clear that a small set is a petite set and a Doeblin set. The next definition is taken from [Ben18].

Definition 1.13. *A point $x^* \in M$ is a* Doeblin point *if there exists a neighbourhood U of x^*, a non trivial measure ξ and $t^* > 0$, such that for all $x \in U$, for all $A \in \mathcal{B}(M)$,*

$$P_{t^*}(x, A) \geq \xi(A).$$

A Doeblin point yields a Doeblin condition provided it is accessible from everywhere in the following sense :

Definition 1.14. *A point $x^* \in M$ is* accessible *from $B \subset M$ if for all neighbourhoods U of x^*, for all $x \in B$, there exists $t > 0$ such that*

$$P_t(x, U) > 0.$$

The next proposition is part of [Ben18, Lemma 4.8] :

Proposition 1.8. *Assume x^* is an accessible Doeblin point and $(P_t)_{t\geq 0}$ is C_b - Feller. Then every compact set is a small set.*

We now give the continuous time counterpart of Theorem 1.2 (see e.g. [Ben18, Theorem 4.10] for a proof) :

Theorem 1.3. *Assume that :*

1. *There exists a function* $V : M \to [0, \infty)$*, and constants* $\gamma \in (0, 1)$*,* $K \geq 0$ *and* $T_0 < T_1$ *such that, for all* $x \in M$*, for all* $t \in [T_0, T_1]$*,*
$$P_t V \leq \gamma V + K.$$
2. *There exists* $R \geq \frac{2K}{1-\gamma}$ *such that the set* $A = \{x \in M : V(x) \leq R\}$ *is a small set.*

Then, $(P_t)_{t \geq 0}$ *admits a unique invariant probability measure* π*. Moreover,* πV *is finite and there exist constants* $\lambda > 0$ *and* $C \geq 0$ *such that, for all* $t \geq 0$*, and every* $f \in L_b(M)$ *with* $\|f\|_V \leq 1$*,*
$$\|P_t f - \pi f\|_V \leq C e^{-\lambda t} \|f - \pi f\|_V.$$
In particular, for all $x \in M$*;*
$$\|\delta_x P_t - \pi\|_V \leq C' e^{-\lambda t}(1 + V(x) + \pi V).$$

From Proposition 1.8, we derive the following corollary of the above theorem :

Corollary 1.1. *Let* $(P_t)_{t \geq 0}$ *be a* C_b *- Feller semigroup such that*

1. *There exists a function* $V : M \to [0, \infty)$*, and constants* $\gamma \in (0, 1)$*,* $K \geq 0$ *and* $T_0 < T_1$ *such that, for all* $x \in M$*, for all* $T \in [T_0, T_1]$*,*
$$P_T V(x) \leq \gamma V(x) + K.$$
2. *For all* $R \geq 0$*, the set* $\{x \in M : V(x) \leq R\}$ *is compact and there exists an accessible Doeblin point.*

Then the conclusions of Theorem 1.3 hold.

Remark 1.4. *When the space* M *is compact, the existence of an accessible Doeblin point is sufficient to have the conclusion of Theorem 1.3. Indeed, the constant function* $V = 0$ *satisfies the first assumption and for every* $R \geq 0$*,* $\{x \in M : V(x) \leq R\} = M$ *which is compact. In that case,* $\|f\|_V = \|f\|$ *and there is uniform exponential convergence in total variation towards the unique invariant probability measure.*

1.2 Piecewise Deterministic Markov Processes

Among the Markov processes, there is one class that is more widely known and studied than any other : Diffusions, or more generally, solutions to Stochastic Differential Equations. However, a large number of phenomena cannot be modelled by diffusions processes. Until the early eighties, no unified theory existed for non-diffusive stochastic processes involving deterministic motions and jumps. Thanks to the work of Davis in 1984 [Dav84] (see also his latter book, [Dav93]), this class of processes has earned its letters of nobility. He baptised them *Piecewise Deterministic Markov Processes* (PDMP), and gave them rigorous basis as a birth present. Let me now introduce to you these young stochastic processes, that were my favourite companions throughout the present thesis.

1.2.1 General Construction of PDMP

A PDMP is a process that evolves deterministically between random times where it jumps to some new random location. Thus the recipe to prepare a good PDMP should contain three ingredients :

- the deterministic way the system moves between the jumps;
- the law of the jump times;
- the law of the location after the jump.

We now give a precise mathematical construction of the PDMP. Let E be a finite set, and M a smooth closed Riemannian submanifold of $\mathbb{R}^d$. We detail the three previous ingredients :

- For each $i \in E$, let $\varphi^i : \mathbb{R} \times M \to M$ be a continuous flow. That is, if we denote $\varphi^i_t : M \to M$, $x \mapsto \varphi^i(t,x)$, then $\varphi^i_0 = Id$ and $\varphi^i_{t+s} = \varphi^i_t \circ \varphi^i_s$ for all $t, s \geq 0$. Whenever ∂M is nonempty, we denote by $t^\star_i(x)$ the first hitting time of the boundary from $(x,i) \in M \times E$:
$$t^\star_i(x) = \inf\{t \geq 0 \ : \varphi^i(t,x) \in \partial M\},$$
with the usual convention $\inf \emptyset = +\infty$.
- For each $i \in E$, we consider a measurable function $\lambda_i : M \to \mathbb{R}_+$ such that, for every $x \in M$, there exists $\varepsilon > 0$ satisfying
$$\int_0^\varepsilon \lambda_i\left(\varphi^i_t(x)\right) dt < \infty.$$
- A Markovian kernel Q on $\mathcal{B}(M \times E)$.

The recipe for the PDMP is as follows. Start at a point $(x,i) \in M \times E$. Pick a random time T_1 with survival function

$$\mathbb{P}_{(x,i)}(T_1 > t) = \begin{cases} \exp\left(-\int_0^t \lambda_i\left(\varphi^i_s(x)\right) ds\right) & \text{if} \quad t \leq t^\star_i(x), \\ 0 \quad \text{if} \quad t > t^\star_i(x). \end{cases}$$

For all $t < T_1$, define $X_t = \varphi_t(x)$ and $I_t = i$. Pick $(y,j) \in M \times E$ with respect to $Q((\varphi_{T_1}(x), i), \cdot)$ and set $X_{T_1} = y$ and $I_{T_1} = j$. Repeat the procedure starting from the new point (y,j) until a random time T_2 selected in a similar fashion. It is easy to believe that the process $(Z_t)_{t\geq 0} = (X_t, I_t)_{t \geq 0}$ that has just been constructed is a Markov process. It is actually even a strong Markov process (see e.g [Dav84, Theorem 25.5]).

From now on, we will work under the following standing assumption, which will be satisfied for all the PDMP considered in this thesis :

Assumption 1.1. *We assume that :*

1. *For all $(x,i) \in M \times E$,*
$$\int_0^{t^\star_i(x)} \lambda_i\left(\varphi^i_t(x)\right) dt = \infty.$$

2. *For $t \geq 0$, let N_t denote the number of jumps occurring before time t. Then for all $(x,i) \in M \times E$ and all $t \geq 0$,*

$$\mathbb{E}_{(x,i)}(N_t) < +\infty.$$

The first assumption guarantees that the process always jumps before reaching the boundary of M, or when there is no boundary, that the probability that the process never jumps is zero. The second assumption prevent an explosion of the number of jumps in finite time.

1.2.2 The infinitesimal generator

The extended generator and its domain $(L, \mathcal{D}(L))$ are given in [Dav84, Theorem 5.5]. Let D_1 be the set of functions $f : M \times E \to \mathbb{R}$, such that for all $i \in E$, for all $x \in M$, $t \mapsto f(\varphi_t^i(x), i)$ is absolutely continuous on $[0, t^\star(x))$, and for $f \in D_1$, set

$$D_{\varphi^i} f(x,i) = \begin{cases} \lim_{t\to 0} \frac{f(\varphi_t^i(x),i) - f(x,i)}{t} & \text{if it exists} \\ 0 \quad \text{otherwise.} \end{cases}$$

Let D_2 be the set of functions $f : M \times E \to \mathbb{R}$ such that there exists a increasing sequence of stopping times σ_n with $\lim_{n\to\infty} \sigma_n = \infty$ and, for all $n \geq 0$,

$$\mathbb{E}_{(x,i)}\left(\sum_{k\geq 0} \mathbb{1}_{T_k < \sigma_n} \left| f(Z_{T_k}) - f(Z_{T_k^-})\right|\right) < +\infty.$$

Then, [Dav84, Theorem 5.5] reads :

Theorem 1.4 (Davis, 1984)**.** *The domain of the extended generator is given by $\mathcal{D}(\bar{L}) = D_1 \cap D_2$. Moreover, for $f \in \mathcal{D}(\bar{L})$, we have*

$$\bar{L}f(x,i) = D_{\varphi^i} f(x,i) + \lambda_i(x) \int_{M\times E} \left(f(y,j) - f(x,i)\right) Q((x,i), dydj). \tag{1.4}$$

By Lemma 1.1, $\mathcal{D}(L) \subset D(\bar{L})$, and for all $f \in \mathcal{D}(L)$, $Lf = \bar{L}f$. Recently, Durmus, Guillin and Monmarché [DGM18] give a general condition for $C_c^1(M)$ to be a core for the strong generator.

1.2.3 A few examples

Since their introduction by Davis, PDMP have become ubiquitous in stochastic modelling of various phenomena. They are applied to neuroscience [PTW10a], [PTW10b], [PTW12], genetics [CDMR12], biology, ecology [BL16], internet traffic [CMP10], [FGM12], [FGM16], [BCG+13]... Recently, their interest for simulations has emerged, see [BRZ17], [BD17], [FBPR18], [BBCD+18], [Mon16] . See also [ABG+14] and [CDG+17] and the references therein for more details and applications. In this section, we give three examples easy to describe and nevertheless giving rise to interesting results.

TCP Process

This first example is called the TCP process, for Transmission Control Protocol. It is a data transmission model of the Internet. For a given connection, the maximal number of packets that can be transmitted at each step is a variable W. If the process is able to transmit all these packets, then at the next round W is increased by 1. If not, the protocol detects a congestion (too much data), and at the next round, W is divided by 2. In a scaling limit, this process can be modelled by a PDMP $(X_t)_{t\geq 0}$ on $[0,+\infty)$ described as follows. Between two jumps, X increases linearly with constant speed 1. Jumps occur at a rate proportional to the position, and at a jump time, X is divided by two. In other words, X is a PDMP on $\mathbb{R}_+$ with the following characteristics :

- For all $t, x \in \mathbb{R}_+$, $\varphi_t(x) = t + x$;
- For all $x \in \mathbb{R}_+$, $\lambda(x) = x$;
- For all $x \in \mathbb{R}_+$, $Q(x, dy) = \delta_{x/2}(dy)$.

The infinitesimal generator of the process if thus given by

$$Lf(x) = f'(x) + x\left(f(\frac{x}{2}) - f(x)\right).$$

In [BCG$^+$13], the exponential ergodicity of the process is proven, using coupling techniques.

Theorem 1.5. *[BCG$^+$13, Theorem 1.5]*
There exist some constants $C, \lambda > 0$ such that, for all $t \geq 0$ and for all initial distributions μ, ν,

$$\|\mu P_t - \nu P_t\|_{TV} \leq Ce^{-\lambda t}.$$

In particular, the process admits a unique invariant probability measure π and

$$\|\mu P_t - \pi\|_{TV} \leq Ce^{-\lambda t}.$$

Linear planar switched systems

Give yourself two 2×2 real matrices A_0 and A_1 and assume they are switched at some random exponentially distributed time with parameters $\lambda_0 > 0$ (from A_0 to A_1) and $\lambda_1 > 0$ (from A_1 to A_0). That is, we are interested in the following process. Start from a point $x \in \mathbb{R}^2$ and $i \in \{0, 1\}$ and let T_1 be a random variable with exponential law of parameter λ_i. Then, for all $t < T_1$, set $X_t = e^{tA_i}x$ and $I_t = i$. Then, at time T_1, i switches from i to $1-i$, so that we have $I_{T_1} = 1 - i$. The component X shall be continuous, so we set $X_{T_1} = e^{T_1 A_i}x$. Then pick a random variable T_2 with exponential law of parameter λ_{1-i}, and for $t \in [T_1, T_1 + T_2[$, set $X_t = e^{(t-T_1)A_{1-i}}X_{T_1}$ and $I_t = 1 - i$, and so on. The process $Z_t = (X_t, I_t)$ is then a PDMP. A natural question is now what is the long-term behaviour of the process X_t ? One could imagine that if, for both A_0 and A_1, all eigenvalues have negative real part, then X_t converges to 0 exponentially fast. Conversely, one could think that if both A_0 and A_1 have an eigenvalue with positive real part, then $\|X_t\|$ will explode. However, this is not the case in general, and the behaviour of X_t depends not only on the individual behaviour of each matrix, but is also highly related to the switching rates λ_0 and λ_1. The two following examples, taken from the concomitant papers by Benaïm, Le Borgne, Malrieu and Zitt [BLBMZ14] and Lawley, Mattingly and Reed [LMR14], exhibit two matrices A_0 and A_1 such that for both of them all the eigenvalues have negative real part, but still, for some switching rate, $\|X_t\|$ goes to infinity. It can be seen as the probabilistic counterpart to the work of Balde, Boscain and Mason [BBM09] on planar system with deterministic switching. A matrix whose all eigenvalues have negative real part is called *Hurwitz*.

Example 1.1. *This example is taken from [BLBMZ14] . Consider the two following matrices :*

$$A_0 = \begin{pmatrix} -1 & 4 \\ 0 & -1 \end{pmatrix}; \quad A_1 = \begin{pmatrix} -1 & 0 \\ 4 & -1 \end{pmatrix}.$$

Then both A_0 and A_1 are Hurwitz. But when looking at $A_{1/2} = \frac{1}{2}A_0 + \frac{1}{2}A_1$, one finds -3 and 1 as eigenvalues. Now consider the switching system as described above, with switching rates $\lambda_0 = \lambda_1 = \beta$, where β is some positive parameter. The intuition is the following. When β is small, only few jumps occur and the switched system follows each individual dynamic for a time long enough to come closer and closer to 0. On the other hand, when β is large, there are many jumps, and the behaviour of X_t is comparable to that of the averaged system, that is the process $Y_t = e^{A_{1/2}t}x_0$. But since $A_{1/2}$ has a positive eigenvalue, we know that $\|Y_t\|$ explodes for almost all initial conditions. Due to the randomness, the same actually happens for $\|X_t\|$ for all nonzero initial conditions. We now quote the precise statement from [BLBMZ14].

Theorem 1.6. *[BLBMZ14, Theorem 1.6] There exist $0 < \beta_1 \leq \beta_2$ such that*

1. *If $\beta < \beta_1$, then X_t converges to 0 exponentially fast for all initial conditions;*
2. *If $\beta > \beta_2$, then $\|X_t\|$ goes to infinity exponentially fast for all nonzero initial conditions.*

Example 1.2. *This second example has been developed in [LMR14]. We have seen in the previous example that one mechanism to obtain an unstable switching from two Hurwitz matrices is the existence of a matrix in the convex hull of the two first ones with a positive eigenvalue. However, it is actually possible to find examples of two matrices such that all convex combinations of them give a Hurwitz matrix and still, for a suitable switching rate, the switched process explodes. This time, the switching rate β has to be chosen neither too small nor too big. Indeed, as explained in the above example, for fast switching, the process is close to the averaged process, which this time converges to 0. Consider the two matrices given by :*

$$A_0 = \begin{pmatrix} -1 & 4 \\ 0 & -1 \end{pmatrix}; \quad A_1 = \begin{pmatrix} -1 & 0 \\ -4 & -1 \end{pmatrix}.$$

It is easily checked that A_0, A_1 and every convex combination of them are Hurwitz. However, one has the following result [LMR14] :

Theorem 1.7. *[LMR14, Example 3.1] There exit $0 < a < b$ such that*

1. *If $\beta \notin (a, b)$, then X_t converges to 0 exponentially fast for all initial conditions;*
2. *There exists $\beta \in (a, b)$ such that $\|X_t\|$ goes to infinity exponentially fast for all nonzero initial conditions.*

Lotka-Volterra in fluctuating environments

This example is taken from [BL16]. In this paper, Benaïm and Lobry study the behaviour of a Lotka-Volterra model, where two species are in competition, in an environment evolving randomly between two states. They arrive at the surprising result that, even if the two states of the environment are, individually taken, favourable to the same species (and thus unfavourable to the other), it can happen that, when switching between these two environments, the disadvantaged species coexist with the favoured one, or even gets the upper hand on it. The results have been generalized to every type of environments by Malrieu and Phu in [MP16]. The model is as follows. Let $E = \{0, 1\}$, and for all $i \in E$, define the vector field F^i on $\mathbb{R}^2$ by

$$F^i(x, y) = \begin{cases} \alpha_i x(1 - a_i x - b_i y) \\ \beta_i y(1 - c_i x - d_i y), \end{cases}$$

where $\alpha_i, \beta_i, a_i, b_i, c_i, d_i$ are positive constants constituting environment $\mathcal{E}_i$. We consider the process $Z = (X, Y, I)$, where $(I_t)_{t\geq 0}$ is a continuous-time Markov chain on E with transition rates λ_0, λ_1; and $(X_t, Y_t)_{t\geq 0}$ is a continuous process with values in $\mathbb{R}^2$, and evolving between the jumps of I as

$$(\dot{X}_t, \dot{Y}_t) = F^{I_t}(X_t, Y_t).$$

We denote by φ^i the flow associated to F^i, that is, for all $(x, y) \in \mathbb{R}_+^2$, $(\varphi_t^i(x, y))_{t\geq 0}$ is the solution of the differential equation $\dot{\zeta} = F^i(\zeta)$ with initial condition $\zeta_0 = (x, y)$. Then Z is a PDMP with flows φ^i, constant jump functions λ_i, and the following Markov kernel Q acting only on E and independently from (x, y) : for all $(x, y, i) \in \mathbb{R}_+^2 \times E$ and all $A \in \mathcal{B}(\mathbb{R}_+^2)$, $Q((x, y, i); A \times \{j\}) = \delta_{(x,y)}(A)\delta_{i\neq j}$. In [BL16], the authors show that there exists a compact set $M \subset \mathbb{R}_+^2$ which is forward invariant for the two vector fields and attracts every solution. In other words, for each $i \in E$, all $(x, y) \in \mathbb{R}_+^2$, there exists a time $T \geq 0$ such that, for every $t \geq T$, $\varphi_t^i(x, y) \in M$. Thus, there is no loss of generality in assuming that the state space of the process Z is $M \times E$. In each fixed environment, the behaviour of the system is easy to describe, and there are only four possibilities :

- If $a < c$ and $b < d$, then for every initial condition $(x, y) \in M$, (x_t, y_t) converges to $(\frac{1}{a}, 0)$: the environment is favourable to species x.
- If $a > c$, and $b > d$, then (x_t, y_t) converges to $(0, \frac{1}{d})$: the environment is favourable to species y.
- If $a > c$ and $b < d$, then there exists a point $(x^*, y^*) \in \mathbb{R}_{++}^2$ such that (x_t, y_t) converges to (x^*, y^*) : this is a coexistence regime.
- If $a < c$ and $b > d$, then there exists an unstable equilibrium point $(x^*, y^*) \in \mathbb{R}_{++}^2$ and (x_t, y_t) converges either to $(\frac{1}{a}, 0)$; $(0, \frac{1}{d})$ or (x^*, y^*) according to the location of the initial condition. This is a bi-stable regime.

One could think that if we choose two environments $\mathcal{E}_1$, $\mathcal{E}_2$ favourable to species x, then if the environment fluctuates between these two states, it is still favourable. However, as announced at the beginning of the section, this is not true in general. In [BL16], the authors show that the long-term behaviour of the process depends on the signs of some quantities called *average growth rates* (see Section 1.3.2 below for details). They show that, depending on these signs, the following four behaviour are possible :

- Almost sure extinction of species y and persistence of x;
- Almost sure extinction of species x and persistence of y;
- Almost sure extinction of at least one species;
- Persistence of the two species (coexistence)

The last case is mathematically translated into the existence of an invariant probability measure Π for Z such that $\Pi(\mathbb{R}^2_{++} \times E) = 1$. The study in [BL16] was done for two environments favourable to the species x. In [MP16], for each pair of environments, examples are given where the resulting PDMP can have the four aforementioned behaviours depending on the jump rates λ_0 and λ_1. More precise statements are given in Section 1.3.2.

1.2.4 Randomly switched vector fields

The Lotka-Volterra model described in the previous section is an example of a particular class of PDMP sometimes referred to "Randomly switched vector fields" or "Vector fields with Markovian switching" (or even "telegraph noise"). These are the kind of PDMPs' that are studied in my thesis, therefore I will now give a more complete description of them, and state some general results, essentially taken from the article *Qualitative properties of certain Piecewise Deterministic Markov Processes* of Benaïm, Malrieu, Le Borgne and Zitt [BLBMZ15].

Construction

Let $d \geq 1$, $E = \{1, \dots, N\}$ a finite set and for all $i \in E$, $F^i : \mathbb{R}^d \to \mathbb{R}^d$ a C^2 globally integrable vector field. We denote by φ^i the flow induced by F^i and we assume that there exists a closed set $M \subset \mathbb{R}^d$ which is forward invariant for all the vector fields, that is

$$\varphi^i_t(M) \subset M, \quad \forall t \geq 0.$$

For all $x \in M$, we are given an irreducible rate matrix $(a_{ij}(x))_{i,j \in E}$. We consider the PDMP $(Z_t)_{t\geq 0} = (X_t, I_t)_{t \geq 0} \in M \times E$, where X is a continuous process on M evolving between the jumps of I according to

$$\frac{dX_t}{dt} = F^{I_t}(X_t), \tag{1.5}$$

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

Set $\lambda_i(x) = \sum_j a_{ij}(x)$ and, for $\lambda_i(x) \neq 0$, set

$$\hat{Q}(x, i, j) = \frac{a_{ij}(x)}{\lambda_i(x)}.$$

The flows φ^i, the rate functions λ_i and the matrix $\hat{Q}$ give the characteristics of the PDMP.

We make the following assumption :

Assumption 1.2. *The jump rates are continuous and bounded :*

$$\sup_{x\in M} \max_{i,j} a_{ij}(x) < \infty.$$

We present now the alternative construction of the process Z made in [BLBMZ15]. By Assumption 1.2, there exists $\lambda > 0$ such that

$$\sup_{x\in M} \max_{i} \lambda_i(x) < \lambda.$$

Set

$$Q(x,i,j) = \frac{a_{ij}(x)}{\lambda} \quad \text{if } i \neq j \quad \text{and} \quad Q(x,i,i) = 1 - \sum_{j\neq i} Q(x,i,j).$$

We first construct a Markov chain that gives the position of the process just after the jumps. Let $(N_t)_{t\geq 0}$ be a Poisson process with parameter λ, and let $(U_n)_{n\geq 0}$ and $(T_n)_{n\geq 0}$ denote the sequences of interjump times and of jump times, respectively. Let $\tilde{Z}_0 = (\tilde{X}_0, \tilde{Y}_0)$ be a random variable independent from $(N_t)_{t\geq 0}$ and construct $\tilde{Z}_n$ recursively by setting :

$$\tilde{X}_{n+1} = \varphi^{\tilde{Y}_n}_{U_{n+1}}(\tilde{X}_n),$$

and by picking $\tilde{Y}_{n+1} \in E$ with the rule

$$\mathbb{P}\left(\tilde{Y}_{n+1} = j \middle| \tilde{X}_{n+1}, \tilde{Y}_n = i\right) = Q(\tilde{X}_{n+1}, i, j).$$

The process $Z_t = (X_t, I_t)$ is constructed by interpolating between the points of the Markov chain :

$$\forall t \in [T_n, T_{n+1}), \quad Z_t = \left(\varphi^{\tilde{Y}_n}_{t-T_n}(\tilde{X}_n), \tilde{Y}_n\right).$$

Inspired by the thinning methods for Poisson processes [LS79], this construction enables to simulate more easily the PDMP since the jumps arrive at exponential times with constant parameter. In particular, one does not need to compute the integral of λ_i along the trajectories of the flow φ^i. The following proposition is [BLBMZ15, Proposition 2.1].

Proposition 1.9. *The semigroup $(P_t)_{t\geq 0}$ is Feller and the strong generator L of the process is given, for every $g \in C^1_c(M \times E)$, by*

$$Lg(x,i) = \langle F^i(x), \nabla g^i(x)\rangle + \lambda \sum_j Q(x,i,j)\left(g(x,j) - g(x,i)\right).$$

Remark 1.5. In [BLBMZ15], the set M is assumed to be compact. However, it is not hard to check that all the arguments given in their proof of Propostion 2.1 go through under Assumption 1.2

Support

We describe now the support of the law of the paths of $(X_t)_{t\geq 0}$. For an initial condition $x \in M$, we write $\mathcal{L}(X^x)$ the law of $(X_t)_{t\geq 0}$ seen as a random variable in $C^0(M)$. For $x \in M$, let

$$\mathrm{co}(F)(x) = \left\{\sum_i \alpha_i F^i(x),\ \alpha_i \geq 0,\ \sum_i \alpha_i = 1\right\}$$

denote the set of the convex combinations of the $F^i(x)$. A solution to the differential inclusion

$$\dot{\eta} \in \mathrm{co}(F)(\eta)$$

is an absolutely continuous function $\eta : \mathbb{R} \to M$ such that $\dot{\eta}_t \in \mathrm{co}(F)(\eta_t)$. For $x \in M$, we let S^x be the set of solutions to the differential inclusion with initial condition x.

Lemma 1.5. *Assume M is compact. Then for every $x \in M$, S^x is a nonempty, compact, connected set.*

The following Theorem is the Support Theorem , [BLBMZ15, Theorem 3.4][1]

Theorem 1.8. *Assume M is compact and $X_0 = x$. Then* $\mathrm{supp}(\mathcal{L}(X^x)) = S^x$.

Accessible set

In this section we describe the set of points that are accessible for the PDMP. For $\mathbf{i} = (i_1, \dots, i_m) \in E^m$ and $\mathbf{u} = (u_1, \dots, u_m) \in \mathbb{R}_+^m$, we denote by $\mathbf{\Phi}_{\mathbf{u}}^{\mathbf{i}}$ the composite flow

$$\mathbf{\Phi}_{\mathbf{u}}^{\mathbf{i}} = \varphi_{u_m}^{i_m} \circ \dots \circ \varphi_{u_1}^{i_1}.$$

For $x \in M$ and $t \geq 0$, we denote by $\gamma_t^+(x)$ (resp. $\gamma^+(x)$) the set of points that are reachable from x at time t (resp. at any nonnegative time) with a composite flow:

$$\gamma_t^+(x) = \{\mathbf{\Phi}_{\mathbf{v}}^{\mathbf{i}}(x),\ (\mathbf{i}, \mathbf{v}) \in E^m \times \mathbb{R}_+^m, m \in \mathbb{N}, v_1 + \dots + v_m = t\},$$

$$\gamma^+(x) = \bigcup_{t \geq 0} \gamma_t^+(x).$$

Definition 1.15. *A point $x^* \in M$ is* $\{F^i\}$-accessible *from $B \subset M$ if $x^* \in \cap_{x \in B} \overline{\gamma^+(x)}$.*

Definition 1.15 actually coïncides with the notion of accessibility for the Markov semi-group defined in Section 1.1 (see e.g. [BLBMZ15, Lemma 3.2], or [BCL17, Lemma 3.1]):

Proposition 1.10. *For all $j, k \in E$, the point $(x^*, j) \in M \times E$ is accessible (for $(P_t)_{t \geq 0}$) from $B \times \{k\} \subset M \times E$ if and only if x^* is $\{F^i\}$-accessible from B.*

Therefore, in the sequel, we will say that a point $x^* \in M$ is accessible from $B \subset M$ if it is $\{F^i\}$-accessible from B. We will simply say that x^* is accessible if it is $\{F^i\}$-accessible from M, and we denote by Γ the (possibly empty) set of points that are accessible. That is,

$$\Gamma = \cap_{x \in M} \overline{\gamma^+(x)}.$$

It turns out that when the set M is compact, Γ has some recurrence properties, summarised in the following statement.

Proposition 1.11. *Assume M is compact and $\Gamma \neq \emptyset$. Let $p \in \Gamma$ and U be a neighbourhood of p. Then there exists open sets $\mathcal{O}_1, \dots, \mathcal{O}_k$ covering M and positive numbers $t_1, \dots, t_k, \delta$ such that, for all $x \in \mathcal{O}_l$ and $i, j \in E$,*

$$\mathbb{P}_{(x,i)}(Z_{t_l} \in \Gamma \times \{j\}) \geq \delta.$$

In particular, if Γ has nonempty interior, then for all $(x, i) \in M \times E$,

$$\mathbb{P}_{(x,i)}(\exists t_0 : \forall t \geq t_0, Z_t \in \Gamma \times E) = 1.$$

[1] Also cf the erratum to [BLBMZ15], to appear at Annales de l'Institut Henri Poincaré

The accessible set is closely linked with the support of the invariant probability measures of the PDMP :

Proposition 1.12. *[BLBMZ15, Proposition 3.17] Assume M is compact. Then*

1. *If $\Gamma \neq \emptyset$, then $\Gamma \times E \subset \mathrm{supp}(\mu)$ for all $\mu \in \mathcal{P}_{inv}$. Furthermore, there exists $\mu \in \mathcal{P}_{inv}$ such that $\Gamma \times E = \mathrm{supp}(\mu)$.*
2. *If Γ has nonempty interior, then $\Gamma \times E = \mathrm{supp}(\mu)$ for all $\mu \in \mathcal{P}_{inv}$.*

It may happen that Γ is empty. In that case, relevant sets for the supports of the invariant probability measures are given by invariant control sets, see [BCL17] for details.

Bracket condition

In the context of diffusion processes, there is a well-known condition ensuring smoothness of transition probabilities and uniqueness of the invariant distribution : the so-called Hörmander condition. It is expressed in terms of the Lie algebra generated by the vector fields generating the stochastic differential equation (see e.g. [Nua06]). Recently, Bakhtin and Hurth [BH12], and Benaïm, Le Borgne, Malrieu and Zitt [BLBMZ15] proved independently that a similar condition can be formulated for randomly switched vector fields and yields a Doeblin type condition.

For two smooth vector fields F and G on $\mathbb{R}^d$, let $[F, G]$ denote the Lie bracket of F and G. It is a vector field on $\mathbb{R}^d$ defined as:

$$[F, G] = \mathrm{D}G.F - \mathrm{D}F.G,$$

where DF stands for the differential of F. Set $\mathcal{F}_0 = \{F^i : i \in E\}$ and construct $\mathcal{F}_k$ recursively as :

$$\mathcal{F}_k = \mathcal{F}_{k-1} \cup \{[F^i, V] \ : i \in E, V \in \mathcal{F}_{k-1}\}.$$

For $x \in M$, we let $\mathcal{F}_k(x)$ be the vector space spanned by $\{V(x); V \in \mathcal{F}_k\}$. We construct $\mathcal{G}_k(x)$ in a similar fashion, starting this time from $\mathcal{G}_0 = \{F^i - F^j; i \neq j\}$.

Definition 1.16. *We shall say that a point $x \in M$ satisfies the* weak bracket condition *if there exists k such that $\mathcal{F}_k(x) = \mathbb{R}^d$. It satisfies the* strong bracket condition *if there exists k such that $\mathcal{G}_k(x) = \mathbb{R}^d$.*

Weak bracket and strong bracket conditions are equivalent to Condition B and Condition A in [BH12], respectively. Note that since $\mathcal{G}_k(x)$ is a subspace of $\mathcal{F}_k(x)$, the strong condition implies the weak one. The converse is not true in general, as illustrated in the following example, taken from [BH12].

Example 1.3. *On the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, consider the two constant vector fields F^1 and F^2 given by the two vectors of the canonical basis of $\mathbb{R}^2$. Then, by definition, for every $x \in \mathbb{T}^2$, $\mathcal{F}_0(x)$ spans $\mathbb{T}^2$ so that the weak bracket condition holds at every point. Now it is also clear that all the Lie brackets generated by F^1 and F^2 are zero, thus for all $k \geq 1$, $\mathcal{G}_k(x) = \mathcal{G}_0(x) = \{F^1 - F^2\}$ which does not span $\mathbb{T}^2$. Hence the strong bracket condition holds nowhere.*

The following theorem is part of [BLBMZ15, Theorem 4.4] or [BH12, Theorem 5].

Theorem 1.9. *Let x be a point of M at which the weak bracket condition holds. Then, there exists $m \geq d$, $\mathbf{i} = (i_1, \ldots, i_m) \in E^m$ and $\mathbf{u} = (u_1, \ldots, u_m) \in \mathbb{R}_+^m$ such that the map $\mathbf{v} \to \mathbf{\Phi}_{\mathbf{v}}^{\mathbf{i}}(x)$ is a submersion at $\mathbf{u}$.*

For $s > 0$ and $m \in \mathbb{N}^*$, we set $D_m^s = \{\mathbf{v} \in \mathbb{R}_+^m : v_1 + \ldots + v_m \leq s\}$. The next result is part of [BLBMZ15, Theorem 4.4] or [BH12, Theorem 4].

Theorem 1.10. *Let x be a point of M at which the strong bracket condition holds. Then, there exist $s > 0$, $i_{m+1} \in E$, $\mathbf{i} \in E^m$ and $\mathbf{u} \in \mathbb{R}_+^m$ with $u_1 + \ldots + u_m < s$ such that the map $D_m^s \to \mathbb{R}^d$, $\mathbf{v} \to \varphi_{s-(v_1+\ldots+v_m)}^{i_{m+1}} \circ \mathbf{\Phi}_{\mathbf{v}}^{\mathbf{i}}(x)$ is a submersion at $\mathbf{u}$.*

This theorem combined with the next one shows the interest of the bracket condition.

Theorem 1.11. *([BLBMZ15, Theorem 4.1]) Let x be a point of M, $(\mathbf{i}, \mathbf{u})$ and $s > u_1 + \ldots + u_m$ such that the map $D_m^s \to \mathbb{R}^d$, $\mathbf{v} \to \varphi_{s-(v_1+\ldots+v_m)}^{i_{m+1}} \circ \mathbf{\Phi}_{\mathbf{v}}^{\mathbf{i}}(x)$ is a submersion at $\mathbf{u}$. Then for all $j \in E$, (x, j) is a Doeblin point.*

Due to the general results in Section 1.1, we deduce the following corollary :

Corollary 1.2. *([BLBMZ15, Theorem 4.6]) Assume that M is compact and that there exists an accessible point at which the strong bracket condition holds. Then, the process Z admits a unique invariant probability measure π. Moreover, there exist positive constants C, γ such that for all $t \geq 0$ and for all $(x, i) \in M \times E$,*

$$\|P_t((x, i), \cdot) - \pi\|_{TV} \leq Ce^{-\gamma t}.$$

In addition, π is absolutely continuous with respect to the Lebesgue measure.

Remark 1.6. *Here and throughout the thesis we shall use a slight abuse of language by calling* Lebesgue measure *the product of the restriction of the Lebesgue measure to M and the counting measure on E. The absolutely continuity of π does not immediately follows from the results stated in Section 1.1. It comes from the fact that the nontrivial measure given by the Doeblin condition in Theorem 1.11 is the Lebesgue measure and the invariance of π (see e.g. [BLBMZ15, Theorem 4.1] and [Ben18, Remark 15]).*

When the weak bracket condition holds instead of the strong one, uniqueness and absolute continuity of the invariant probability measure is still guaranteed. This is because the weak bracket condition yields a Doeblin point for the embedded Markov chain $\tilde{Z}$.

Proposition 1.13. *([BLBMZ15, Theorem 4.5]) Assume that there exists an accessible point at which the weak bracket condition holds. Then, the process Z admits a unique invariant probability measure π which is absolutely continuous with respect to Lebesgue measure.*

1.3 Stochastic Persistence

This section is devoted to the presentation of the theory of stochastic persistence. It will be essentially focused on the preprint of Michel Benaïm [Ben18]. When studying a mathematical model for ecology or epidemiology, one of the most natural questions is whether a species will persist or die in the long run. For nonrandom system, this question has been considered for many decades yet. One of the central issue is to determine conditions insuring that a system is persistent or not. In the early eighties, Hofbauer introduced the concept of Lyapunov average functions [Hof81], that would lead to several persistence results [Sch00], [GH03], [HS04]. The goal of the paper [Ben18] is to generalise this notion to stochastic models and to extend results given for Stochastic Differential Equations in [BHS08] and [SBA11]. The general idea goes as follows. Consider a Markov process $(X_t)_{t\geq 0}$ on some metric space (M,d) and assume that X leaves invariant some closed subset $M_0 \subset M$. That is, for all $t \geq 0$, $X_t \in M_0$ if and only if $X_0 \in M_0$. The set M_0 can be seen as the *extinction* set of the process X. Persistence is roughly speaking given by a function V exploding at M_0, such that $\mu H < 0$ for all invariant probability measure μ of X on M_0, where $H = \bar{L}V$ and $\bar{L}$ is the extended generator of X. Replacing $\mu H < 0$ by $\mu H > 0$ gives extinction. We make these statements more precise in the next subsection.

1.3.1 Definitions and results

Let M be a locally compact metric space and X a càdlàg Markov process on M. We assume that $(P_t)_{t\geq 0}$ is Feller. We let L and $\bar{L}$ denote the strong and the extended generator, respectively, of (P_t) and $\mathcal{D}(L)$ and $\mathcal{D}(\bar{L})$ their domains. We let $\mathcal{D}^2 \subset \mathcal{D}(L)$ denote the set of $f \in \mathcal{D}(L)$ such that $f^2 \in \mathcal{D}(L)$. For $f \in \mathcal{D}^2(L)$ the *Carré du champ* of f is defined as

$$\Gamma f = L(f^2) - 2fLf.$$

We assume that

Assumption 1.3. *There exists a nonempty compact set $M_0 \subset M$ called the* extinction set *which is invariant under $(P_t)_{t\geq 0}$. That is*

$$P_t \mathbb{1}_{M_0} = \mathbb{1}_{M_0}, \quad \forall t \geq 0.$$

We set

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

The set M_+ is the set where the process is not extinct. Note that M_+ is also invariant under $(P_t)_{t\geq 0}$. To avoid problems at infinity, we assume that there exists a Lyapunov function. Recall that a proper function $W : M \to \mathbb{R}_+$ is a continuous function such that, for all $R > 0$, the set $\{x \in M : W(x) \leq R\}$ is compact.

Assumption 1.4. *There exist proper maps $W, \tilde{W} : M \to \mathbb{R}_+$, a continuous function $\mathcal{L}W : M \to \mathbb{R}$ and a positive constant C such that*

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

   (a) *$W|_K = W_K|_K$ and $(LW_K)|_K = \mathcal{L}W_K|_K$,*

   (b) *For all $x \in M$,*

   $$\sup\{P_t(\Gamma W_K)(x), t \geq 0,\ K \text{ compact}\} < +\infty$$

2. $$\mathcal{L}W \leq -\tilde{W} + C.$$

Remark 1.7. *As noticed in [Ben18], Assumption 1.4 is automatically satisfied when M is compact with $W = \tilde{W} = \mathcal{L}W = 0$.*

We let
$$\Pi_t = \frac{1}{t}\int_0^t \delta_{X_s} ds$$
denote the empirical occupation measure of X. Assumption 1.4 ensures that the family of random measures $(\Pi_t)_{t\geq 0}$ is almost surely tight and the invariance of weak limit points.

Theorem 1.12. *[Ben18, Theorem 2.1] Assume Assumption 1.4 holds. Then for all $x \in M$, the family $(\Pi_t)_{t\geq 0}$ is $\mathbb{P}_x$-almost surely tight, and every limit point of $(\Pi_t)_{t\geq 0}$ lies in $\mathcal{P}_{inv}(M)$. Moreover, $\mathcal{P}_{inv}(M)$ is compact.*

Extinction of X amounts to say that trajectories of (X_t) converge almost surely to M_0. Let M_0^ε be the ε-neighborhood of M_0. Using a terminology borrowed to Schreiber [Sch12] and Chesson [Che82], we say that X is *stochastically persistent* (or *almost surely persistent*) provided
$$\lim_{\varepsilon\to 0}\limsup_{t\to\infty}\Pi_t(M_0^\varepsilon) = 0$$
$\mathbb{P}_x$ almost surely for all $x \in M_+$. We say that X is *persistent in probability* if
$$\lim_{\varepsilon\to 0}\limsup_{t\to\infty}\mathbb{P}_x(X_t \in M_0^\varepsilon) = 0$$
for all $x \in M_+$. In addition to Assumptions 1.3 and 1.4, we assume the following.

Assumption 1.5. *There exist continuous maps $V : M_+ \to \mathbb{R}^+$ and $H : M \to \mathbb{R}$ enjoying the following properties :*

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

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

(c) $\lim_{x\to M_0} V(x) = \infty$;

(d) *Jumps of $V(X_t)$ are bounded : $\exists\Delta > 0$ such that for all $t \geq 0$, $|V(X_t)-V(X_{t-})| \leq \Delta$;*

(e) *The map $\frac{\tilde{W}}{1+|H|}$ is proper.*

Remark 1.8. *When M is compact, since we can choose $\tilde{W} = 0$, point* **(e)** *is always satisfied.*

It turns out that under this assumption, V is in the domain of the extended generator. We quote [Ben18, Lemma 7.4]:

Proposition 1.14. *For all $x \in M_+$, the process*
$$M_t^V(x) = V(X_t) - V(x) - \int_0^t H(X_s)ds$$
is a $\mathbb{P}_x$ - martingale, square integrable, such that, $\mathbb{P}_x$-almost surely, $\lim_{t\to\infty}\frac{M_t^V(x)}{t} = 0$. In particular, the function V is in the extended domain of the generator (seen as a process on M_+), and $\bar{L}V = H$.

Let $\mathcal{P}_{erg}(M_0) = \mathcal{P}_{erg} \cap \mathcal{P}(M_0)$. Define the H-exponents of the processes as

$$\Lambda^+(H) = -\inf_{\mu \in \mathcal{P}_{erg}(M_0)} \mu H \text{ and } \Lambda^-(H) = -\sup_{\mu \in \mathcal{P}_{erg}(M_0)} \mu H.$$

We call the process *H-persistent* if $\Lambda^-(H) > 0$ and *H-nonpersistent* if $\Lambda^+(H) < 0$.

The following theorem is a consequence of Theorems 4.4 and 4.10 and Proposition 8.2 in [Ben18].

Theorem 1.13. *Suppose that the process X is H-persistent. Then*

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

(ii) *For every T_0 large enough and $T_1 > T_0$, there exist $0 < \rho < 1$ and positive constants θ, K, δ such that for all $x \in M_+$ and $T \in [T_0, T_1]$,*

$$P_T(e^{\theta V})(x) \leq \begin{cases} \rho e^{\theta V}(x) & \text{if} \quad d(x, M_0) < \delta \\ K & \text{otherwise}; \end{cases}$$

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

$$\tau^\varepsilon = \inf\{t \geq 0 : \ X_t \in M_0^\varepsilon\}.$$

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

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

(iv) *If, furthermore, there exists a Doeblin point $x \in M_+$ accessible from M_+ and α such that $\tilde{W} = \alpha W$, then $\mathcal{P}_{inv}(M_+)$ reduces to a single measure Π and for all $x \in M_+$*

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

for some $\kappa, C > 0$.

The next result is a general extinction result.

Theorem 1.14. *Suppose that the process X is H-nonpersistent. Then*

(i) *For all $0 < \alpha < -\Lambda^+(H)$, there exists a neighborhood U of M_0 and $\eta > 0$ such that*

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

for all $x \in U$;

(ii) *If furthermore M_0 is accessible from M,*

$$\mathbb{P}_x(\liminf_{t \to \infty} \frac{V(X_t)}{t} \geq -\Lambda^+(H)) = 1$$

for all $x \in M_+$.

We provide here a sketch of the proof of Theorem 1.14 in the case where M is compact, similar to the one given in [BL16, Theorem 3.1].

Proof Let $0 < \alpha < -\Lambda^+(H)$. The proofs of Propositions 8.2 and 8.3 in [Ben18] (see also [BL16, Lemma 3.5]) adapt verbatim in the nonpersistent case to prove that there exist $T > 0$, $\theta > 0$, $\varepsilon > 0$ and $0 < \rho < 1$ such that, for all $z \in M_0^{V,\varepsilon} \setminus M_0$,

(i) $P_T V(z) - V(z) \geq \alpha T$,

(ii) $P_T e^{-\theta V}(z) \leq \rho e^{-\theta V(z)}$.

Here and throughout this proof, $M_0^{V,\varepsilon} = \{z \in M_+ \ : \ V(z) > -\log(\varepsilon)\} \cup M_0$. We set $\tau_\varepsilon = \inf\{k \geq 0 \ : \ X_{kT} \notin M_0^{V,\varepsilon}\}$. We claim that :

(a) There exists $\eta > 0$ such that for all $x \in M_0^{V,\varepsilon/2}$, $\mathbb{P}_x(\tau_\varepsilon = \infty) \geq \eta$;

(b) On the event $\{\tau_\varepsilon = \infty\}$, and for all $x \in M_0^{V,\varepsilon/2}$, $\liminf_{t\to\infty} \frac{V(X_t)}{t} \geq \alpha$.

In particular, this implies point **(i)** of the theorem with $U = M_0^{V,\varepsilon/2}$. Point **(ii)** easily follows by the Markov property. We prove point **(a)**. We set for $k \geq 0$, $W_k = e^{-\theta V(X_{kT})}$. Due to point (ii) above, $(W_{k\wedge\tau_\varepsilon})_{k\geq 0}$ is a supermartingale. In particular, for all $x \in M_0^{V,\varepsilon/2} \setminus M_0$,

$$\mathbb{E}_x(W_{k\wedge\tau_\varepsilon} \mathbb{1}_{\tau_\varepsilon<\infty}) \leq e^{-\theta V(z)} \leq \left(\frac{\varepsilon}{2}\right)^\theta.$$

By dominated convergence, this gives

$$\varepsilon^\theta \mathbb{P}_x(\tau_\varepsilon < \infty) \leq \left(\frac{\varepsilon}{2}\right)^\theta,$$

which proves the first point with $\eta = 1 - 2^{-\theta}$. We now prove point **(b)**. We set

$$M_n = \sum_{k=1}^{n} \left(P_T V(X_{(k-1)T}) - V(X_{kT})\right).$$

The sequence $(M_n)_{n\geq 1}$ is a martingale, and on the event $\{\tau_\varepsilon = \infty\}$ and for all $x \in M_0^{V,\varepsilon/2} \setminus M_0$,

$$\frac{M_n}{nT} \geq \alpha - \frac{V(X_{nT})}{nT}.$$

Furthermore, the quadratic variation $\langle M \rangle$ of M is given by

$$\langle M \rangle_n = \sum_{k=0}^{n-1} \mathbb{E}_x\left[(P_T V(X_{kT}) - V(X_{(k+1)T}))^2 | \mathcal{F}_{kT}\right].$$

We claim that for all $x \in M_+$, there exists a constant C_x such that, for all $k \geq 0$,

$$\mathbb{E}_x\left[(P_T V(X_{kT}) - V(X_{(k+1)T}))^2\right] \leq (2T\|H\| + \sqrt{C_x T})^2. \tag{1.6}$$

In particular, for all $x \in M_+$, the sequence $(n^{-1}\mathbb{E}_x(\langle M \rangle_n))_{n\geq 1}$ is bounded and the strong law of large numbers for martingales (see e.g [Duf97, Theorem 1.3.17]) implies that $\mathbb{P}_x$ - almost surely,

$$\lim_{n\to\infty} \frac{M_n}{n} = 0.$$

Hence, on the event $\{\tau_\varepsilon = \infty\}$ and for all $x \in M_0^{V,\varepsilon/2} \setminus M_0$,

$$\liminf_{n\to\infty} \frac{V(X_{nT})}{nT} \geq \alpha. \tag{1.7}$$

Now, by Proposition 1.14, $\mathbb{P}_x$ - almost surely,

$$\lim_{t\to\infty} \frac{1}{t}\left(V(X_t) - \int_0^t H(X_s)ds\right) = 0.$$

Since H is bounded, this implies that for all $t \in [0,T]$,

$$\lim_{n\to\infty} \frac{V(X_{nT+t}) - V(X_{nT})}{nT} = 0.$$

This, together with (1.7) proves point **(b)**. It remains to show that (1.6) holds. For convenience, we write $\|\cdot\|_2$ for the L^2 norm under $\mathbb{P}_x$. By triangular inequality,

$$\|P_T V(X_{kT}) - V(X_{(k+1)T}\|_2 \leq \|P_T V(X_{kT}) - V(X_{kT})\|_2 + \|V(X_{kT}) - V(X_{(k+1)T})\|_2.$$

Since for all $y \in M_+$, $M^V(y)$ is a martingale under $\mathbb{P}_y$, $P_T V(y) - V(y) = \int_0^T P_s H(y)ds$. In particular, H being bounded, $\|P_T V - V\| \leq T\|H\|$ and thus

$$\|P_T V(X_{kT}) - V(X_{kT})\|_2 \leq T\|H\|.$$

On the other hand, we have

$$V(X_{(k+1)T}) - V(X_{kT}) = M^V_{(k+1)T}(x) - M^V_{kT}(x) - \int_{kT}^{(k+1)T} H(X_s)ds$$

Still by triangular inequality and boundedness of H, to prove(1.6) it suffices to show that $\|M^V_{(k+1)T}(x) - M^V_{kT}(x)\|_2 \leq \sqrt{C_x T}$. By Proposition 1.14, $M^V(x)$ is a square integrable martingale; thus

$$\mathbb{E}_x\left((M^V_{(k+1)T}(x) - M^V_{kT}(x))^2\right) = \mathbb{E}_x\left(\langle M^V(x)\rangle_{(k+1)T} - \langle M^V(x)\rangle_{kT}\right)$$

Using a stopping time argument and Hypothesis 1.5 **(b)**, and reasoning as in the proof of [Ben18, Lemma 7.4], we can show that

$$\mathbb{E}_x\left(\langle M^V(x)\rangle_{(k+1)T} - \langle M^V(x)\rangle_{kT}\right) \leq C_x T,$$

where

$$C_x = \sup\{P_t(\Gamma V_K)(x), t \geq 0,\ K \subset M_+, K \,\text{compact}\} < +\infty.$$

This proves the claim.

□

Remark 1.9. *When M_0 is noncompact, things become trickier and one should add a condition of "persistence at infinity" (see [Ben18, Section 8.1]). Since for all the models studied in this thesis, M_0 is compact, we won't detail here the noncompact case.*

1.3.2 An example : PDMP Lotka-Volterra Model

Here we shall apply the above theory to the example of the competitive Lotka-Volterra model with switching introduced in Section 1.2.3 and exhaustively studied in [BL16] and [MP16]. Recall that this is a PDMP $Z_t = (X_t, Y_t, I_t)$ where $(I_t)_{t\geq 0}$ is a continuous-time Markov chain on $E := \{0, 1\}$ with transition rate λ_0, λ_1; and $(X_t, Y_t)_{t\geq 0}$ is a continuous process with values in $\mathbb{R}^2$, and evolving between the jumps of I as

$$(\dot{X}_t, \dot{Y}_t) = F^{I_t}(X_t, Y_t),$$

with

$$F^i(x, y) = \begin{cases} \alpha_i x(1 - a_i x - b_i y) \\ \beta_i y(1 - c_i x - d_i y). \end{cases}$$

One can check that for $\eta > 0$ small enough, the compact set

$$K = \{(x, y) \in \mathbb{R}^2_+ \ : \eta \leq x + y \leq 1/\eta\}$$

is positively invariant for the flows φ^i and attracts every solution starting from $\mathbb{R}^2_+ \setminus \{0\}$. Hence we shall consider $M := K \times E$ as the state space of the process Z. For this model, there are two extinction sets : $M_0^{\mathrm{x}} = K_0^{\mathrm{x}} \times E$ and $M_0^{\mathrm{y}} = K_0^{\mathrm{y}} \times E$, corresponding respectively to the extinction of species x and y. Writing $K_0^{\mathrm{y}} = \{(x, y) \in K \ : y = 0\}$, it is straightforward that the strong bracket condition holds at every point of K_0^{y} for dynamics restricted to K_0^{y}. Moreover, for the process restricted to K_0^{y}, the accessible set is the interval $[p_-, p_+]$, where $p_- = \min(\frac{1}{a_0}, \frac{1}{a_1})$ and $p_- = \max(\frac{1}{a_0}, \frac{1}{a_1})$. Thus by Corollary 1.2, the process Z restricted to M_0^{y} admits a unique invariant probability measure μ. The same reasoning proves that Z admits a unique invariant probability measure $\hat{\mu}$ on M_0^{x}. We can now construct a Lyapunov function V and test the criteria for persistence. Let $V : \mathbb{R}^*_+ \to \mathbb{R}_+$ be a smooth function coinciding with $-\log$ on $(0, 1]$, and for $(x, y, i) \in M_+^{\mathrm{x}}$ define $V^{\mathrm{x}}(x, y, i) = V(x)$ and for $(x, y, i) \in M_+^{\mathrm{y}}$, $V^{\mathrm{y}}(x, y, i) = V(y)$. For $(x, y, i) \in M$, define $H^{\mathrm{y}}(x, y, i) = -\beta_i(1 - c_i x - d_i y)yV'(y)$ and $H^{\mathrm{x}}(x, y, i) = -\alpha_i(1 - a_i x - c_i y)xV'(x)$, where we set $xV'(x)|_{x=0} = 1$. We have the following lemma :

Lemma 1.6. *The functions* $(V^{\mathrm{y}}, H^{\mathrm{y}})$ *and* $(V^{\mathrm{x}}, H^{\mathrm{x}})$ *satisfy Assumption 1.5.*

Proof We only prove the lemma for $(V^{\mathrm{y}}, H^{\mathrm{y}})$. For all $\varepsilon > 0$, let $V_\varepsilon^{\mathrm{y}}$ be a bounded smooth function coinciding with V^{y} on $\{y > \varepsilon\}$. By Proposition 1.9, the generator L of Z acts on $V_\varepsilon^{\mathrm{y}}$ as

$$LV_\varepsilon^{\mathrm{y}}(x, y, i) = \langle F^i(x, y), \nabla V_\varepsilon^{\mathrm{y}}(x, y, i)\rangle + \lambda_i \left(V_\varepsilon^{\mathrm{y}}(x, y, 1 - i) - V_\varepsilon^{\mathrm{y}}(x, y, i)\right),$$

which gives

$$LV_\varepsilon^{\mathrm{y}}(x, y, i) = -\beta_i y(1 - c_i x - d_i y)\partial_y V_\varepsilon^{\mathrm{y}}(x, y, i).$$

Since $V_\varepsilon^{\mathrm{y}}(x, y, i) = -V(y)$ for all $y > \varepsilon$, we get that $LV_\varepsilon^{\mathrm{y}}(x, y, i) = H^y(x, y, i)$ for all (x, y, i) such that $y > \varepsilon$. This proves point **(a)** of Assumption 1.5. Points **(b)**, **(c)** and **(d)** are immediate from the definitions, and point **(e)** follows from Remark 1.8. □

Note that $H^{\mathrm{y}}(x, 0, i) = -\beta_i(1-c_i x)$. Since $\mathcal{P}_{erg}(M_0^{\mathrm{y}}) = \{\mu\}$, there is only one H-exponents on M_0^{y}. We denote it by Λ_{y} :

$$\Lambda_{\mathrm{y}} = \Lambda^+(H^{\mathrm{y}}) = \Lambda^-(H^{\mathrm{y}}) = \int_{M_0^{\mathrm{y}}} \beta_i(1 - c_i x)\mu(dx, i).$$

Similarly, we denote by Λ_{x} the average growth rate of x :

$$\Lambda_{\mathrm{x}} = \Lambda^{+}(H^{\mathrm{x}}) = \Lambda^{-}(H^{\mathrm{x}}) = \int_{M_0^{\mathrm{x}}} \alpha_i(1 - b_i y)\hat{\mu}(dy, i).$$

For the sake of simplicity, we assume now that both environments $\mathcal{E}_0$ and $\mathcal{E}_1$ are favourable to species x, as in [BL16]. Thanks to a result of Malrieu and Zitt [MZ17], we have the following lemma :

Lemma 1.7. *If $\Lambda_{\mathrm{y}} > 0$ and $\Lambda_{\mathrm{x}} < 0$, then M_0^{x} is accessible.*

The results described in the previous section can now be applied to prove the following theorem, which sums up [BL16, Theorems 3.1, 3.3, 3.4, 4.1]

Theorem 1.15. *There are four possible regimes :*

1. *If $\Lambda_{\mathrm{y}} < 0$ and $\Lambda_{\mathrm{x}} > 0$, then for all $(x, y, i) \in M \setminus M_0^{\mathrm{x}}$,*

$$\mathbb{P}_{(x,y,i)}\left(\limsup_{t\to\infty} \frac{1}{t}\log(Y_t) \leq \Lambda_{\mathrm{y}}\right) = 1.$$

2. *If $\Lambda_{\mathrm{y}} > 0$ and $\Lambda_{\mathrm{x}} < 0$, then for all $(x, y, i) \in M \setminus M_0^{\mathrm{y}}$,*

$$\mathbb{P}_{(x,y,i)}\left(\limsup_{t\to\infty} \frac{1}{t}\log(X_t) \leq \Lambda_{\mathrm{x}}\right) = 1.$$

3. *Denote by $\mathrm{Ext}_{\mathrm{y}}$ and $\mathrm{Ext}_{\mathrm{x}}$ the events described in the above probabilities. If $\Lambda_{\mathrm{y}} < 0$ and $\Lambda_{\mathrm{x}} < 0$, then for all $(x, y, i) \in M$,*

$$\mathbb{P}_{(x,y,i)}\left(\mathrm{Ext}_{\mathrm{x}}\right) + \mathbb{P}_{(x,y,i)}\left(\mathrm{Ext}_{\mathrm{y}}\right) = 1$$

   and

$$\mathbb{P}_{(x,y,i)}\left(\mathrm{Ext}_{\mathrm{y}}\right) > 0$$

   if $(x, y, i) \notin M_0^{\mathrm{x}}$.

4. *If $\Lambda_{\mathrm{y}} > 0$ and $\Lambda_{\mathrm{x}} > 0$ then the process admits a unique invariant probability measure Π such that $\Pi(M_+) = 1$. Moreover,*

   (a) Π is absolutely continuous with respect to the Lebesgue measure on $K \times E$,

   (b) there exists $\theta > 0$ such that

$$\int_{M_+} \left(\frac{1}{x^\theta} + \frac{1}{y^\theta}\right)\Pi(dx, dy, i) < \infty,$$

   (c) For all $(x, y, i) \in M_+$, $\lim_{t\to\infty} \Pi_t = \Pi$, $\mathbb{P}_{(x,y,i)}$-almost surely,

   (d) There exists $C, \lambda > 0$, such that, for all $(x, y, i) \in M_+$,

$$\|P_t((x, y, i), \cdot) - \Pi\| \leq C(1 + \frac{1}{x^\theta} + \frac{1}{y^\theta})e^{-\lambda t}.$$

Remark 1.10. *In [BL16], the last point of Theorem 1.15 is not proven when $\frac{\beta_0\alpha_1}{\alpha_0\beta_1} = \frac{a_0c_1}{c_0a_1} = \frac{b_0d_1}{d_0b_1}$. This case is a consequence of a result of this thesis, detailed in Chapter 2.*

1.4 Random Dynamical Systems and Lyapunov exponents

Some results of the present thesis rely on the links that one can make between randomly switched vector fields and Random Dynamical Systems. The latter may be seen as the stochastic extension of dynamical systems and (semi)flows. The theory of Random Dynamical Systems has received a lot of attention in the eighties and nineties through the work of Arnold and his co-authors. The reader is referred to the book of Arnold [Arn98] for more information.

1.4.1 Definitions

Let $\mathbb{T}$ be one of the following (semi)groups of time : $\mathbb{R}, \mathbb{R}_+, \mathbb{N}, \mathbb{Z}$. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a semigroup of transformations of Ω, $(\theta_t)_{t\in\mathbb{T}}$, preserving the measure $\mathbb{P}$. That is, for all $t \in \mathbb{T}$, $\mathbb{P} \circ \theta_t^{-1} = \mathbb{P}$. We say that $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t\in\mathbb{T}})$ is a *metric Dynamical System* (DS).

Definition 1.17. *Let $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t\in\mathbb{T}})$ be a metric dynamical system. A set $A \subset \Omega$ is said to be* invariant *if $\theta_t^{-1}A = A$ for all $t \in \mathbb{T}$. If all invariant sets have probability* 0 *or* 1*, then the dynamical system is said to be* ergodic.

With a slight abuse of notation, we will say that θ is ergodic if the metric DS is ergodic. A classical example of metric DS is given by a stationary Markov Chain. Let $(X_n)_{n\geq 0}$ be a canonical Markov chain on some measurable state space E . That is, $\Omega = E^{\mathbb{N}}$, $\mathcal{F}$ is the Borel sigma algebra of Ω and $(X_n)_{n\geq0}$ is the canonical process, i.e for all $n \geq 0$, $X_n : \Omega \to E$ is defined for all $(\omega_n)_{n\geq0}$ by $X_n((\omega_n)_{n\geq0}) = \omega_n$. For all $x \in E$, $\mathbb{P}_x$ is a probability on Ω such that $(X_n)_{n\geq0}$ is a Markov Chain under the family $(\mathbb{P}_x)_{x\in E}$ and $\mathbb{P}_x(X_0 = x) = 1$. Consider the shift operator $\theta : \Omega \to \Omega$ defined for all $\omega = (\omega_n)_{n\geq0} \in \Omega$ by $\theta(\omega) = (\omega_{n+1})_{n\geq0}$. Now assume that $(X_n)_{n\geq0}$ admits an invariant probability measure p on E. Then $(\Omega, \mathcal{F}, \mathbb{P}_p, (\theta^n)_{n\geq0})$ is a metric DS (see e.g [DMPS18]).

Let $(\mathcal{X}, \mathcal{B})$ be a measurable space.

Definition 1.18. *Let $\varphi : \mathbb{T} \times \Omega \times \mathcal{X} \to \mathcal{X}$ be a measurable map. We say that (φ, θ) is a* Random Dynamical System *(RDS) if :*

1. *For all $(\omega, x) \in \Omega \times \mathcal{X}$,*
$$\varphi(0, \omega, x) = x.$$

2. *For all $t, s \in \mathbb{T}$, for all $(\omega, x) \in \Omega \times \mathcal{X}$,*
$$\varphi(t + s, \omega, x) = \varphi(t, \theta_s\omega, \varphi(s, \omega, x)). \tag{1.8}$$

Equation (1.8) is referred to the *cocycle property*, hence the alternative name of *cocycle* for Random Dynamical System. It is the random counterpart of the semiflow property in a deterministic dynamical system.

Definition 1.19. *We give the following definitions*

1. *We say that a RDS (φ, θ) is* ergodic *if θ is ergodic.*

2. *We say that a RDS is linear if for all $(t, \omega) \in \mathbb{T} \times \Omega$, the map $\varphi(t, \omega, \cdot)$ is linear.*

1.4.2 Generation of RDS

Product of random matrices

For a given integer $d \geq 1$, we denote by $M_d(\mathbb{R})$ the set of real $d \times d$ matrices. Let $A : \Omega \to M_d(\mathbb{R})$ be a measurable map, and for $n \geq 0$, set $A_n(\omega) = A(\theta^n \omega)$. Consider the discrete dynamical system given by

$$x_{n+1} = A_n x_n.$$

Starting at some point x_0, it satisfies $x_n = A_{n-1} \cdots A_0 x_0$. This is the canonical example of a discrete-time RDS. By the cocycle property (1.8), every linear discrete-time RDS (φ, θ) can be represented in this way by setting $A(\omega) = \varphi(1, \omega)$. It entails in particular systems of the form

$$x_{n+1} = A(\xi_{n+1}) x_n,$$

where $(\xi_n)_{n\geq 0}$ is a stationary Markov chain. This kind of processes have been the focus of much attention : [FK60], [Bou87], [Bou88], [AGD94], [AC97]...

Continuous-time RDS

Let $\mathbb{T} = \mathbb{R}_+$, $\mathcal{X} = \mathbb{R}^d$ and $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t\geq 0})$ be a metric dynamical system. We are interested in random differential equations of the form

$$\dot{x}_t = f(\theta_t \omega, x_t),$$

for some function $f : \Omega \times \mathbb{R}^d \to \mathbb{R}^d$. Such a system is called a Random Differential Equation. The following results, quoting Theorems 2.2.1 and 2.2.13 in [Arn98], state that there is equivalence between RDS and Random Differential Equations :

Theorem 1.16. *With the above notation, we have the following :*

1. *If for all ω and $t \in \mathbb{R}$, $x \mapsto f(\theta_t \omega, x)$ is locally Lipschitz and if for all $K \subset \mathbb{R}^d$ compact, $t \mapsto \|f(\theta_t \omega, \cdot)\|_{BL(K)}$ is locally integrable, then*

$$\dot{x}_t = f(\theta_t \omega, x_t)$$

admits a unique maximal solution $\varphi(t, \omega, x)$ which is a RDS over θ. Here for $g : \mathbb{R}^d \to \mathbb{R}^d$ locally Lipschitz, $\|g\|_{BL(K)}$ stands for

$$\|g\|_{BL(K)} = \sup_{x\in K} \|g(x)\| + \sup_{x,y\in K, x\neq y} \frac{\|g(x) - g(y)\|}{\|x - y\|}.$$

2. *Let φ be a RDS such that for all (ω, x), the map $t \mapsto \varphi(t, \omega, x)$ is differentiable at $t = 0$. Set*

$$f(\omega, x) = \frac{d}{dt} \varphi(t, \omega, x)|_{t=0}.$$

Then f is measurable and

$$\frac{d}{dt} \varphi(t, \omega, x) = f(\theta_t \omega, \varphi(t, \omega, x)).$$

PDMP as RDS

In this section we describe a PDMP obtained from randomly switched vector fields with constant jump rates as a RDS. That is, let $(I_t)_{t\geq 0}$ be a continuous-time Markov chain on some finite state space E. For each $i \in E$, let $F^i : \mathbb{R}^d \to \mathbb{R}^d$ be a globally integrable vector field, and let X_t be the solution of

$$\frac{dX_t}{dt} = F^{I_t}(X_t).$$

Assume that $(I_t)_{t\geq 0}$ is irreducible on E. Thus, there exists a unique invariant probability measure p for I, which is therefore ergodic. As in the discrete-time setting, we can construct $(I_t)_{t\geq 0}$ as the canonical process on the set Ω of càdlàg functions from $\mathbb{R}_+$ to E endowed with its Borel sigma algebra $\mathcal{F}$, and supporting a family of probability measures $(\mathbb{P}_x)_{x\in E}$ such that $\mathbb{P}_x(I_0 = x) = 1$. Let θ_t be the usual shift operator on $\Omega :$ for all $\omega \in \Omega$,

$$\theta_t\omega(s) = \omega(t+s).$$

Then $(\Omega, \mathcal{F}, \mathbb{P}_p, (\theta_t)_{t\geq 0})$ is an ergodic metric dynamical system, where $\mathbb{P}_p = \sum_{x\in E} p_x \mathbb{P}_x$. In this setting, one can rewrite

$$\frac{dX_t}{dt} = f(\theta_t\omega, X_t),$$

where $f(\omega, x) = F^{\omega(0)}(x)$ for all $x \in \mathbb{R}^d$ and all $\omega = (\omega_t)_{t\geq 0} \in \Omega$. Thus (X, θ) is an ergodic RDS.

1.4.3 Multiplicative Ergodic Theorem and its consequence

The most fundamental result in the theory of Random Dynamical System concerns linear RDS : the famous Oseledets Mutliplicative Ergodic Theorem (see [Arn98, Theorem 3.4.1]). It gives a nice spectral theory for linear random differential equation. According to Arnold, one can count no less than fifteen different proofs of this theorem since its first publication by Oseledets in the sixties [Ose68]. We consider a continuous time linear RDS. According to Theorem 1.16, there exists a measurable map $A : \Omega \to M_d(\mathbb{R})$ such that

$$\frac{d\varphi(t,\omega,x)}{dt} = A(\theta_t\omega)\varphi(t,\omega,x).$$

We write Φ for the fundamental matrix of φ. That is, for all $(t, \omega, x) \in \mathbb{R}_+ \times \Omega \times \mathbb{R}^d$, $\varphi(t,\omega,x) = \Phi(t,\omega)x$. We will identify φ with Φ.

Theorem 1.17 (Multiplicative Ergodic Theorem, Oseledets). *Let $\Phi(t,\omega)$ be a linear RDS such that $\Phi(t,\omega) \in GL_d(\mathbb{R})$ for almost all $(t,\omega) \in \mathbb{R}_+ \times \Omega$. Assume furthermore that*

$$\mathbb{E}\left[\sup_{t\in[0,1]} \log^+ \|\Phi(t,\omega)\|\right], \mathbb{E}\left[\sup_{t\in[0,1]} \log^+ \|\Phi(t,\omega)^{-1}\|\right] < \infty. \tag{1.9}$$

Then there exists a subset $\tilde{\Omega}$ of Ω, of full measure and such that $\theta_t\tilde{\Omega} = \tilde{\Omega}$ for all $t \geq 0$ on which the following hold :

1. *There exists* $q(\omega) \in \{1, \dots d\}$*, numbers* $\lambda_{q(\omega)}(\omega) < \dots < \lambda_1(\omega)$ *and a sequence of linear subspaces*
$$\{0\} = V_{q(\omega)+1}(\omega) \subset V_{q(\omega)}(\omega) \subset \cdots \subset V_1(\omega) = \mathbb{R}^d$$
such that
$$\lim_{t\to\infty} \frac{1}{t} \log \|\varphi(t, \omega, x)\| = \lambda_i(\omega)$$
if and only if $x \in V_i(\omega) \setminus V_{i+1}(\omega)$.

2. *The maps* $\omega \mapsto q(\omega)$, $\omega \mapsto \lambda_i(\omega)$, $\omega \mapsto V_i(\omega)$ *and* $\omega \mapsto d_i(\omega)$ *are measurable. Here* $d_i(\omega)$ *stands for the codimension of* $V_{i+1}(\omega)$ *in* $V_i(\omega)$.

3. *For all* $t \geq 0$, $q(\theta_t\omega) = q(\omega)$, $\lambda_i(\theta_t\omega) = \lambda_i(\omega)$ *and* $\Phi(t, \omega)V_i(\omega) = V_i(\theta_t\omega)$.

4. *If* $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t\geq 0})$ *is ergodic, then* q, λ_i *and* d_i *are constant functions on* $\tilde{\Omega}$.

The functions (or numbers in the ergodic case) λ_i are called the *Lyapunov exponents* of (Φ, θ) (or of A). The collection of the Lyapunov exponents together with their multiplicities is called the Lyapunov Spectrum of (Φ, θ). The maximal Lyapunov exponent λ_1 is also called the principal Lyapunov exponent. It satisfies, for all $\omega \in \tilde{\Omega}$,
$$\lim_{t\to\infty} \frac{1}{t} \log \|\Phi(t, \omega)\| = \lambda_1(\omega).$$
(see also Kingman's subadditive ergodic theorem, [Arn98, Theorem 3.3.2]).

Remark 1.11. *One can also give a discrete time version of the MET, holding for discrete time RDS. Details can be found in [Arn98, Theorem 3.4.1].*

We quote the following corollary [Arn98, Corollary 3.3.4].

Corollary 1.3. *Under the assumptions of the above theorem, for all* $\omega \in \tilde{\Omega}$,
$$\lim_{t\to\infty} \frac{1}{t} \log |\det \Phi(t, \omega)| = \sum_{i=1}^{p(\omega)} d_i(\omega)\lambda_i(\omega).$$

For the particular case where the matrices $A(\omega)$ are block triangular, a result from Hennion [Hen84, Proposition 1] implies that the Lyapunov spectrum of A coincides with the union of the Lyapunov spectrum of the matrices on the diagonal :

Theorem 1.18. *Assume that for some* m, n *such that* $m + n = d$*, there exist measurable maps* $B : \Omega \to M_n(\mathbb{R})$, $C : \Omega \to M_{m,n}(\mathbb{R})$ *and* $D : \Omega \to M_m(\mathbb{R})$ *such that for all* ω,
$$A(\omega) = \begin{pmatrix} B(\omega) & 0 \\ C(\omega) & D(\omega) \end{pmatrix}.$$
Suppose that the integrability condition (1.9) *is satisfied and let* $\mathcal{S}(A)$, $\mathcal{S}(B)$, $\mathcal{S}(D)$ *denote the set of Lyapunov exponents counted with multiplicity of* A, B *and* D*, respectively. Then*
$$\mathcal{S}(A) = \mathcal{S}(B) \cup \mathcal{S}(D).$$

Remark 1.12. Proposition 1 in [Hen84] is given for a discrete-time random dynamical system. However, its proof adapts verbatim to the continuous-time case by using the continuous-time version of the Multiplicative Ergodic Theorem given above. One could also have used [GMO08, Theorem 1.1].

Application to PDMP

We have already proven that a linear PDMP with constant jump rates can be represented as a RDS. If furthermore the vector fields F^i are linear, then the resulting RDS is also linear. Under this assumption, we denote by A^i the matrix such that $F^i(x) = A^i x$. Now we show that the condition (1.9) is satisfied for $\Phi(t,\omega)$ which is the solution to the matrix equation

$$\frac{d\Phi(t,\omega)}{dt} = A^{\theta_t\omega}\Phi(t,\omega),$$

with initial condition $\Phi(0,\omega) = Id$. We also have that

$$\frac{d\Phi(t,\omega)^{-1}}{dt} = -\Phi(t,\omega)^{-1}A^{\theta_t\omega}.$$

In particular, if $K = \max_i \|A^i\|$, it follows from Gronwall's lemma that

$$\|\Phi(t,\omega)\|, \|\Phi(t,\omega)^{-1}\| \leq e^{Kt}$$

for all $t \geq 0$. This implies that condition (1.9) is satisfied, whence the conclusion of Theorem 1.17 hold for the PDMP, with constant q, λ_i and d_i. In particular, we have proved the following statement.

Proposition 1.15. *Consider a finite family of $d \times d$ matrices $(A^i)_{i\in E}$, the set Ω of càdlàg functions on E and $\mathbb{P}_p$ a probability measure on Ω such that the canonical process $(I_t)_{t\geq 0}$ is a stationary Markov Chain on E. For $\omega \in \Omega$ and $x \in \mathbb{R}^d$, let $\varphi(t,\omega,x)$ denote the solution of*

$$\frac{d\varphi(t,\omega,x)}{dt} = A^{I_t(\omega)}\varphi(t,\omega,x)$$

such that $\varphi(0,\omega,x) = x$. Then there exists a number $q \in \{1,\dots,d\}$, numbers $\lambda_1 > \ldots > \lambda_q$ and $\tilde{\Omega} \subset \Omega$ with $\mathbb{P}_p(\tilde{\Omega}) = 1$ such that, for all $\omega \in \tilde{\Omega}$, there is a sequence of linear subspaces

$$\{0\} = V_{q+1}(\omega) \subset V_q(\omega) \subset \cdots \subset V_1(\omega) = \mathbb{R}^d$$

such that

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

if and only if $x \in V_i(\omega) \setminus V_{i+1}(\omega)$.

Simple principal Lyapunov exponent

In some situation, it is possible to have further information on the Lyapunov spectrum of the RDS. An interesting question is for example under which conditions the dominating Lyapunov exponent λ_1 is simple (that is $d_1 = 1$). For the case of products of i.i.d random 2×2 matrices a well-known theorem of Furstenberg gives such a condition (see e.g. [BL85, Theorem 4.4]). For products of positive random matrices, a random version of the Perron Frobenius Theorem has been given in [AGD94]. In the particular case where $\det \Phi(t,\omega) = 1$, knowing that $d_1 = 1$ ensures that the principal Lyapunov exponent is positive. In [AC97], the authors showed that the set of invertible matrices with simple Lyapunov spectrum is dense in the set of invertible matrices. Here we detail a more general result from Bougerol [Bou88] applying to multiplicative systems, that will be used in Chapter 4.

Let $(\sigma_t)_{t\geq 0}$ be a stationary Markov process on some metric space $\mathcal{E}$, and $(M_t)_{t\geq 0}$ a process with values in $GL_d(\mathbb{R})$. We introduce the following definition, which is [Bou88, Definition 1.1 and 1.2].

Definition 1.20. *Let π be a probability measure on $\mathcal{E}$. We say that (M,σ,π) is a multiplicative system, if :*

(i) *The process (M,σ) is Markovian with semigroup $(P_t)_{t\geq 0}$;*

(ii) *For any Borel subset $A\subset\mathcal{E}$ (resp. $B\subset GL_d(\mathbb{R})$), $t\geq 0$, $C\in GL_d(\mathbb{R})$ and $x\in E$, one has*
$$P_t\left((C,x);BC\times A\right)=P_t\left((\mathrm{Id},x);B\times A\right),$$
where $BC=\{NC;N\in B\}$;

(iii) *π is an ergodic measure for σ and* $\sup_{0\leq t\leq 1}\mathbb{E}_{\mathrm{Id},\pi}\left[\log^+\|M_t\|+\log^+\|M_t^{-1}\|\right]<\infty$.

We denote by U the first-order resolvent of $(P_t)_{t\geq 0}$, which is defined via
$$U=\int_0^{+\infty}\mathrm{e}^{-t}P_t\mathrm{d}t.$$
For $x\in\mathcal{E}$ let D_x be the support of $U((\mathrm{Id},x),\cdot)$ and $S_x=\{B\in GL_d(\mathbb{R})\ :\ (B,x)\in D_x\}$. We will also let K be the first order resolvent of the semigroup $(R_t)_{t\geq 0}$ of the Markov process σ.

Definition 1.21. *We say that a multiplicative system (M,σ,π) satisfies hypothesis **H** if the following conditions hold*

(i) *The space $\mathcal{E}$ is a complete metric space.*

(ii) *The semigroup $(P_t)_{t\geq 0}$ is Feller.*

(iii) *The support of π is $\mathcal{E}$. If h is a bounded measurable function which is a fixed point for the first order resolvent of σ, i.e.*
$$Kh=h,$$
then h is continuous.

One has the following result (see [Bou88, Theorem 1.7]).

Theorem 1.19 (Bougerol, 1988). *Assume (M,σ,π) is a multiplicative system satisfying hypothesis **H**. Assume furthermore that*

(i) *For some $x\in\mathcal{E}$, there exists a matrix in S_x with a simple eigenvalue of biggest modulus.*

(ii) *There does not exist some finite union W of proper vector spaces of $\mathbb{R}^d$ such that, for all matrices M in S_x, $MW=W$.*

Then $d_1=1$.

Remark 1.13. *Theorem 1.19 is a reformulation of [Bou88, Theorem 1.7], which is given for the numbers γ_i that are the Lyapunov exponents for the external power of M (see [Bou88, Proposition 2.2] or [Arn98, Theorem 3.3.3] for details). The numbers γ_i are the numbers λ_i counted with multiplicity (see [Arn98, Definition 3.3.8 and Theorem 3.4.1]).*

In Chapter 4, we show that a linear PDMP is a multiplicative system, and use Theorem 1.19 to prove a conjecture by Takeushi et al. on a Lotka-Volterra prey-predator system with random switching.

1.5 Quasi-stationary distribution

Consider a Markov process $(X_t)_{t\geq 0}$ living on some metric space (M,d) containing an extinction set. That is, a closed subset M_0 of M, positively invariant under X : for all $t,s\geq 0$, $X_t\in M_0 \Rightarrow X_{t+s}\in M_0$. Then, essentially three long term-behaviours can be observed : extinction in finite time, extinction in infinite time or persistence. The two last cases are the ones considered in Section 1.3. In the persistence situation, the process admits an invariant distribution giving no mass to M_0. This is of course not possible if extinction occurs in finite time : in this situation, every invariant probability measure concentrates on the extinction set. However, it is interesting to describe the behaviour of the process before the extinction time, and in many situations, a metastable equilibrium of the process can arise. This is what is called of a *Quasi-Stationary Distribution*, whose definition is given now. We consider a càdlàg Markov process X such that for all $x\in M_+ = M\setminus M_0$,

$$\mathbb{P}_x(T_0 < +\infty) = 1,$$

where $T_0 = \inf\{t\geq 0 \;:\; X_t\in M_0\}$. We also assume that for all $x\in M_+$ and for all $t>0$,

$$\mathbb{P}_x(t < T_0) > 0,$$

that is, the probability to survive at least until time t is positive for every t.

Definition 1.22. *A probability measure α on M_+ is a* Quasi-Stationary Distribution *(QSD) if for all $t\geq 0$, for all $A\in\mathcal{B}(M_+)$,*

$$\mathbb{P}_\alpha(X_t\in A | t<T_0) = \alpha(A). \tag{1.10}$$

The reader interested in a survey on QSD is advised to read the excellent one by Méléard and Villemonais [MV12]. We recall the following classical result on QSD (see e.g. [MV12, Proposition 2]).

Proposition 1.16. *Let α be a QSD. Then there exists $\lambda>0$ such that, for all $t\geq 0$,*

$$\mathbb{P}_\alpha(T_0 > t) = e^{-\lambda t}.$$

When looking at the process X, the *killed semigroup* is of particular interest. It is the family of operators $(\tilde{P}_t)_{t\geq 0}$ defined as :

$$\tilde{P}_t f(x) = \mathbb{E}_x\left(f(X_t)\mathbb{1}_{t<T_0}\right).$$

In particular, the above proposition reads

$$\alpha\tilde{P}_t f = e^{-\lambda t}\alpha f.$$

We let $(\tilde{L}, \mathcal{D}(\tilde{L}))$ denote the infinitesimal generator associated to the killed semigroup. That is, the operator as defined in Section 1.1 with P_t replaced by $\tilde{P}_t$. The next proposition is the analogue of Proposition 1.3 (see e.g. [MV12, Proposition 4]) :

Proposition 1.17. *Assume that there exists a set $D \subset \mathcal{D}(\tilde{L})$ such that, for all $A \in \mathcal{B}(M_+)$, there exists a uniformly bounded sequence of functions f_n in D that converges pointwise to $\mathbb{1}_A$. Let α be a probability measure on M_+. Then α is a QSD if and only if there exists $\lambda > 0$ such that for all $f \in D$, for all $t \geq 0$,*

$$\alpha\tilde{L}\tilde{P}_t f = -\lambda\alpha\tilde{P}_t f.$$

In that case, λ coincides with the one given by Proposition 1.16.

Remark 1.14. *In [MV12], the assumption given in Proposition 4 is that $\alpha\tilde{L}f = -\lambda\alpha f$ for all $f \in D$. However, it not sufficient, and the stronger assumption on $\tilde{P}_t f$ we give is implicitly used in their proof.*

Finding conditions ensuring the existence and uniqueness of a QSD is not always easy. Recently, in an impressive series of articles [CV16], [CV17a], [CV17d], [CV17c], [CV17b], [CV18], [CCPV18], Nicolas Champagnat and Denis Villemonais developped a general theory for existence and uniqueness of QSD relying on Lyapunov and Doeblin - type conditions. We first present the necessary and sufficient condition given in [CV16] for uniform exponential convergence to the QSD. It is called Assumption (A) in [CV16].

Assumption (A) There exists a probability measure ν on M_+ such that

(A1) There exist $t_0, c_1 > 0$ such that for all $x \in M_+$,

$$\mathbb{P}_x(X_{t_0} \in \cdot | t < T_0) \geq c_1\nu(\cdot).$$

(A2) There exists $c_2 > 0$ such that for all $x \in M_+$ and all $t \geq 0$,

$$\mathbb{P}_\nu(t < T_0) \geq c_2\mathbb{P}_x(t < T_0).$$

We can now quote the main theorem of [CV16] :

Theorem 1.20. *[CV16, Theorem 2.1] Assumption (A) holds if and only if there exists a unique QSD α and constants $C, \gamma > 0$ such that, for all $\mu \in \mathcal{P}(M_+)$, for all $t \geq 0$,*

$$\|\mathbb{P}_\mu(X_t \in \cdot | t < T_0) - \alpha\|_{TV} \leq Ce^{-\gamma t}.$$

Moreover, if Assumption (A) holds, one can choose $C = 2(1 - c_1c_2)$ and $\gamma = -\frac{\log(1-c_1c_2)}{t_0}$.

This theorem can be seen as the absorbed counterpart of Theorem 1.1. We refer to [CV16, Remark 2] for the main difference between these two results. As in the non-absorbed case, it can be very difficult to obtain (A1) when the state space M_+ is noncompact. In [CV17b], the authors give a Lyapunov - type criterion to still get exponential convergence, as in Theorem 1.3 above. The condition is given both in discrete and continuous time. Although this thesis is concerned with continuous-time processes, we give the discrete-time version that will be used in Chapter 6. Let P denote the submarkovian kernel of an almost surely absorbed Markov chain $(X_n)_{n\geq 0}$:

$$Pf(x) = \mathbb{E}_x(f(X_1)\mathbb{1}_{1<T_0}).$$

Condition (E) There exist a measurable subset $K \subset M_+$, a probability measure ν on K, integers $n_1, n_2 \geq 1$, positive constants $c_1, c_2, c_3, \theta_1, \theta_2$, functions $\varphi_1, \varphi_2 : M_+ \to \mathbb{R}_+$ such that

(E1) For all $x \in K$,
$$\mathbb{P}_x(X_{n_1} \in \cdot) \geq c_1 \nu(\cdot \cap K).$$

(E2) $\theta_1 < \theta_2 \leq 1$ and :

(a) $\inf_{x \in M_+} \varphi_1(x) \geq 1$ and $\sup_{x \in K} \varphi_1(x) < +\infty$,

(b) $\inf_{x \in K} \varphi_2(x) > 0$ and $\sup_{x \in M_+} \varphi_2(x) \leq 1$,

(c) For all $x \in M_+$,
$$P\varphi_1(x) \leq \theta_1 \varphi_1(x) + c_2 \mathbb{1}_K(x)$$
and
$$P\varphi_2(x) \geq \theta_2 \varphi_2(x).$$

(E3)
$$\sup_{n \geq 1} \frac{\sup_{x \in K} \mathbb{P}_x(n < T_0)}{\inf_{x \in K} \mathbb{P}_x(n < T_0)} \leq c_3.$$

(E4) For all $x \in K$, there exists $n_4(x)$ such that for all $n \geq n_4(x)$,
$$\mathbb{P}_x(X_n \in K) > 0.$$

We now quote the main result in [CV17b] :

Theorem 1.21. *[CV17b, Theorem 2.1] Under Condition* **(E)***, there exist $C > 0$, $\rho \in (0,1)$ and a probability measure α such that, for all $\mu \in \mathcal{P}(M_+)$ with $\mu\varphi_1 < +\infty$ and $\mu\varphi_2 > 0$,*
$$\|\mathbb{P}_\mu(X_n \in \cdot | n < T_0) - \alpha\|_{TV} \leq C\rho^n \frac{\mu\varphi_1}{\mu\varphi_2}.$$
Moreover, α is the unique QSD of X satisfying $\alpha\varphi_1 < +\infty$ and $\alpha\varphi_2 > 0$, and there exist $\delta > 0$ depending only on φ_1, φ_2, θ_1, θ_2, and c_2 such that $\alpha(K) \geq \delta$.

1.6 What is done in this thesis

To conclude, we sum up the main original results that are obtained in this thesis.

1.6.1 Random Switching between Vector Fields sharing a common zero

Based on a joint work with Michel Benaïm [BS19] :

> M. Benaïm and E. Strickler, *Random Switching between Vector Fields having a common zero*, Annals of Applied Probability, **29** (2019), no. 1, 326-375.

This work can be found in Chapter 3 and deals with randomly switched system having a common equilibrium point. That is, we consider a PDMP $(Z_t)_{t \geq 0} = (X_t, I_t)_{t \geq 0}$ on $\mathbb{R}^d \times E$ defined as in Section 1.2.4 by a family of finitely many globally integrable vector fields $F^i : \mathbb{R}^d \to \mathbb{R}^d$ and rate functions $(a_{i,j}(x))_{i,j \in E}$ satisfying the following assumption :

Assumption 1.6. *For all $i \in E$, $F^i(0) = 0$.*

The first natural examples coming to mind of vector fields satisfying this assumption are linear vector fields. Looking at Examples 1.1 and 1.2, even in simple cases, the behaviour of the process may be complicated, and depend highly on the switching rates. Nonetheless, it is still easier to study than randomly switched non-linear vector fields. The goal of our work was to give a probabilistic counterpart to the following well-known deterministic result :

Theorem 1.22. *Let $F : \mathbb{R}^d \to \mathbb{R}^d$ be a globally integrable smooth vector field, and let x_t be the solution of $\dot{x}_t = F(x_t)$. Assume that $F(0) = 0$ and set $A = DF(0)$ the Jacobian matrix of F at 0. Finally, let $\lambda(A)$ be the largest real part of the eigenvalues of A. Then*

1. *If $\lambda(A) < 0$, there exists a neighbourhood of 0 such that if x_0 belongs to this neighbourhood, then x_t converges to 0 exponentially fast;*

2. *If $\lambda(A) > 0$, then 0 is unstable : there exists a neighbourhood of 0 such that one can find x_0 arbitrary close to 0 for which the solution starting at x_0 leaves this neighbourhood.*

We learned from Examples 1.1 and 1.2 that looking at the eigenvalues of each Jacobian matrix and of their convex combinations is not enough to come to similar conclusion for the randomly switched system. The right object to look at is the average growth rate of the linear system. Note that in the deterministic case, if y_t denotes the solution of $\dot{y}_t = Ay_t$ with initial condition $y_0 \neq 0$, then

$$\lim_{t\to\infty} \frac{1}{t} \log \|y_t\| \leq \lambda(A),$$

with equality for almost all starting points. In the random case, let Y_t be the solution of $\dot{Y}_t = A^{J_t} Y_t$, where for all $i \in E$, $A^i = DF^i(0)$ and $(J_t)_{t\geq 0}$ is the continuous-time Markov chain on E with transition rates $(a_{i,j}(0))_{i,j\in E}$. Whenever $Y_0 \neq 0$, one can consider the polar decomposition of Y_t. That is, we set $\Theta_t = \frac{Y_t}{\|Y_t\|}$ and $\rho_t = \|Y_t\|$ to obtain

$$\begin{cases} \frac{d\Theta_t}{dt} = A^{J_t}\Theta_t - \langle A^{J_t}\Theta_t, \Theta_t\rangle\Theta_t \\ \frac{d\rho_t}{dt} = \rho_t\langle A^{J_t}\Theta_t, \Theta_t\rangle. \end{cases} \tag{1.11}$$

Remark 1.15. *For stochastic differential equations, the idea of introducing this polar decomposition goes back to Hasminskii [Has60] and has proved to be a fundamental tool for analysing linear stochastic differential equations (see e.g [Bax91]), linear random dynamical systems (see e.g chapter 6 of Arnold [Arn98]) and more recently certain linear PDMPs in [BLBMZ14], [LMR14] or [Lag16].*

Several interesting consequences can be drawn from Equation (1.11). Firstly, the process (ρ, Θ, J) is still a PDMP, and so is also the process (Θ, J) since the evolution of Θ does not depend on ρ. Secondly, the exponential growth rate of Y can be expressed as follows :

$$\frac{1}{t}\log\|Y_t\| = \frac{1}{t}\log\rho_t = \frac{1}{t}\int_0^t \langle A^{J_s}\Theta_s, \Theta_s\rangle ds + \frac{\log\rho_0}{t}.$$

Now, if μ is an ergodic invariant probability measure of (Θ, J), then Birkhoff's Ergodic Theorem implies that for μ - almost every point $(\theta_0, i_0) \in S^{d-1} \times E$, one has

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \langle A^{J_s}\Theta_s, \Theta_s\rangle ds = \int_{S^{d-1}\times E} \langle A^i\theta, \theta\rangle d\mu(\theta, i), \quad \mathbb{P}_{(\theta_0,i_0)} - \text{a.s.}$$

In particular, for μ - almost every point $(\theta_0, i_0) \in S^{d-1} \times E$, one has

$$\lim_{t\to\infty} \frac{1}{t} \log \|Y_t\| = \int_{S^{d-1}\times E} \langle A^i\theta, \theta\rangle d\mu(\theta, i), \quad \mathbb{P}_{(\theta_0,i_0)} - \text{a.s.}$$

This motivates the definition of the following quantities, that we call *average growth rates* : for all invariant probability measures μ of (Θ, J),

$$\Lambda(\mu) = \sum_{i\in E} \int_{S^{d-1}} \langle A^i\theta, \theta\rangle \mu^i(d\theta), \tag{1.12}$$

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

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

These quantities play the same role as the eigenvalues in the deterministic case. That is why we are interested in their extremal values :

$$\Lambda^- = \inf\{\Lambda(\mu) \ : \mu \in \mathcal{P}_{erg}^{(\Theta,J)}\} \text{ and } \Lambda^+ = \sup\{\Lambda(\mu) \ : \mu \in \mathcal{P}_{erg}^{(\Theta,J)}\}. \tag{1.13}$$

By Feller continuity of (Θ, J) and compactness of $S^{d-1} \times E$, this infimum and supremum are actually minimum and maximum. The sign of these quantities will determine the long-term behaviour of the non linear process (X, I). Before stating the precise result, we give a theorem linking the average growth rate to the Lyapunov exponents as defined in Section 1.4.3. Recall that the PDMP $(Y_t)_{t\geq 0}$ can be seen as a Random Dynamical System satisfying the integrability condition of the Multiplicative Ergodic Theorem. We denote by $\lambda_1 > \ldots > \lambda_p$ the associated Lyapunov exponents given in Proposition 1.15. The following statement is part of Proposition 3.1.

Proposition 1.18. *For all ergodic probability measures μ of (Θ, J), one has*

$$\Lambda(\mu) \in \{\lambda_1, \ldots, \lambda_p\}.$$

Moreover, λ_1 is attained : $\Lambda^+ = \lambda_1$.

In general, there is no reason to think that λ_p can be obtained as an average growth rate. A reason for this is that the random vector space associated to λ_p given in the Multiplicative Ergodic Theorem is measurable with respect to the future. For the matrices given in Example 1.1, it can be shown that the process (Θ, J) is uniquely ergodic, and therefore $\Lambda^+ = \Lambda^-$. Moreover, we know from Theorem 1.6 that $\Lambda^+ > 0$ provided the switching rate is sufficiently large. On the other hand, it can be deduced from Corollary 1.3 that

$$\sum_i d_i\lambda_i = \frac{1}{2}\left(\mathrm{Tr}(A^0) + \mathrm{Tr}(A^1)\right) < 0.$$

Since $\lambda_1 = \Lambda^+ > 0$, this shows that $d_1 = d_2 = 1$ and $\lambda_2 < 0$. In particular, $\Lambda(\mu) > \lambda_2$ for the unique invariant probability measure of the process (Θ, J).

The nonlinear process

We now state the main results obtained in Chapter 3. For the sake of simplicity, we assume that there exists a compact set $M \subset \mathbb{R}^d$ containing 0 which is forward invariant for all the flows φ^i, and thus for the PDMP. Therefore, we shall consider $M \times E$ as the state space of Z. Setting $M_0 = \{0\}$ and $M^* = M_+ = M \setminus \{0\}$, the sets $M_0 \times E$ and $M^* \times E$ are invariant for Z. The first result is an *extinction* result, which can be viewed as the probabilistic counterpart of the first point of Theorem 1.22. It is part of Theorem 3.1.

Theorem 1.23. *Assume $\Lambda^+ < 0$. Let $0 < \alpha < -\Lambda^+$. Then there exists a neighbourhood $\mathcal{U}$ of 0 (in M) and $\eta > 0$ such that for all $x \in \mathcal{U}$ and $i \in E$*

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

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

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

The next result is a *persistence* result obtained under the assumption that $\Lambda^- > 0$. It represents the probabilistic counterpart of the second point of Theorem 1.22 and is part of Theorem 3.2. We recall that Π_t is the empirical distribution measure of Z defined as

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

Theorem 1.24. *Assume $\Lambda^- > 0$. Then the following assertions hold:*

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

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

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

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

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

(iii) *Let $\varepsilon > 0$ and let 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 $i \in E$,

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

Thanks to the general results on PDMP presented in Section 1.2.4, one can obtain the following theorems. The first one ensures the uniqueness of the invariant probability measure giving full mass to $M^* \times E$ under the weak bracket condition. It also provides an ergodic theorem.

Theorem 1.25. *In addition to the assumption $\Lambda^- > 0$, assume that there exists a point $p \in M^*$, F^i-accessible from M^*, at which the weak bracket condition holds. Then*

(i) *The set $\mathcal{P}^Z_{inv} \cap \mathcal{P}(M^* \times E)$ reduces to a single element, denoted Π;*

(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.

The second theorem states that, under a stronger condition the process converges to its unique invariant probability measure on $M^* \times E$.

Theorem 1.26. *Under the conditions of the preceding theorem, assume furthermore that one of the two following holds :*

(i) *The weak bracket condition is strengthened to the strong bracket condition; or*

(ii) *There exist $\alpha_1, \dots, \alpha_N \in \mathbb{R}$ with $\sum \alpha_i = 1$ and a point $e^\star \in M^*$ accessible from M^* such that $\sum \alpha_i F^i(e^\star) = 0$.*

Then there exist $\kappa, \theta, C > 0$ such that for all $x \in M^*$ and $i \in E$,

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

The two latter theorems are Theorem 3.3 and 3.4, respectively.

Application to SIS model in random environment

We give an example of application of the previous results to the SIS epidemiological model considered in the Introduction. That is, we consider a vector field F of the form (2). Lajmanovich and Yorke [LY76] prove the following result:

Theorem 1.27 (Lajmanovich and Yorke, [LY76])**.** *Let $A = C - \mathsf{diag}(D)$.*

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

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

We now consider a set of matrices C^i and of vectors D^i, and we let F^i denote the vector field induced by C^i and D^i as in equation (2). We consider a continuous-time irreducible Markov chain on E, $(I_t)_{t \geq 0}$, and the PDMP given by

$$\frac{dX_t}{dt} = F^{I_t}(X_t). \tag{1.14}$$

Thanks to the above results, we are able to prove the following theorem, which can be seen as a probabilistic counterpart to Theorem 1.27. It is part of Theorems 3.9 and 3.10.

Theorem 1.28. *Denote by Λ^+ and Λ^- the extremal growth rates induced by (1.14). Then*

1. *$\Lambda^+ = \Lambda^-$;*

2. *If $\Lambda^+ < 0$, then for all $(x,i) \in [0,1]^d \times E$,*

$$\mathbb{P}_{(x,i)}\left(\lim_{t\to\infty} \frac{1}{t}\log\|X_t\| \leq \Lambda^+\right) = 1;$$

3. *If $\Lambda^+ > 0$, then the process $Z = (X, I)$ admits a unique invariant probability measure Π on $[0,1]^d \setminus \{0\} \times E$. Furthermore, there exists a distance-like function $\tilde{d}$ and $r > 0$, such that, for all $x \neq 0$, for all $t \geq 0$,*

$$\mathcal{W}_{\tilde{d}}(\delta_{x,i}P_t, \Pi) \leq e^{-rt}\mathcal{W}_{\tilde{d}}(\delta_{x,i}, \Pi).$$

Here $\mathcal{W}_{\tilde{d}}$ stands for the Wasserstein distance induced by $\tilde{d}$.

In the case where $\Lambda^+ > 0$, note that we do not require the existence of an accessible point satisfying a bracket condition to have exponential convergence towards the invariant probability measure. This is because monotonicity of the vector fields provide a contraction for the semigroup $(P_t)_{t\geq 0}$ (see Section 3.6 in Chapter 3).

1.6.2 Application to a PDMP Lotka-Volterra prey-predator model

Based on a joint work with Alexandru Hening [HS19] :

A. Hening and E. Strickler, *On a predator-prey system with random switching that never converges to its equilibrium*, accepted at SIAM Journal on Mathematical Analysis (2019).

This work can be found in Chapter 4 and describe a result obtained on a Lotka-Volterra prey-predator model with switching. In the paper [HS19], we answer a conjecture raised by Takeuchi et al. in [TDHS06]. With the above notations, we consider two vector fields F^0 and F^1 on $\mathbb{R}^2$ given by

$$F^i(x,y) = \begin{pmatrix} x(a_i - b_i y) \\ y(-c_i + d_i x) \end{pmatrix}. \tag{1.15}$$

Note that the vector field F^i has a unique positive equilibrium $(p_i, q_i) = (c_i/d_i, a_i/b_i)$.The random switching between two such vector fields has been studied in [TDHS06]. In that paper, the authors distinguish two cases :

I. $p_0 = p_1 =: p$ and $q_0 = q_1 =: q$, i.e. common zero for F^0 and F^1,

II. $(p_0, q_0) \neq (p_1, q_1)$, i.e. different zeroes for F^0 and F^1.

In case II, they are able to show that the process oscillates between zero and infinity and leaves any compact subset of $\mathbb{R}^2_{++} := \{(x,y) \in \mathbb{R}^2_+ \; xy > 0\}$. However, in case I, they are not able to determine whether there is the same long-term behaviour as in case II or whether there is convergence of the system to the common equilibrium (p,q) :

Theorem 1.29 (Takeuchi et al., 2006)**.** *For any $(x_0, y_0) \in \mathbb{R}^2_{++}$ and $i \in \{0,1\}$, either*

$$\lim_{t\to\infty} X_t = (p,q), \quad \mathbb{P}_{(x_0,y_0,i)} - \text{a.s.} \tag{1.16}$$

or

$$\limsup x_t = \limsup y_t = +\infty, \quad \liminf x_t = \liminf y_t = 0, \quad \mathbb{P}_{(x_0,y_0,i)} - \text{a.s.} \tag{1.17}$$

It was conjectured from simulations (see [TDHS06, Remark 5.1]) that only case 1.17 happens in the above theorem. We prove that this is indeed the case by showing that the unique average growth rate $\Lambda = \Lambda^+ = \Lambda^-$ of the process at point (p, q) is positive. In particular, by Theorem 1.24 above, case 1.16 cannot happen and thus only case 1.17 occurs. Note that with the notation $A^i = DF^i(p, q)$, one has :

$$A^i = \begin{pmatrix} 0 & -\alpha_i \\ \beta_i & 0 \end{pmatrix},$$

where $\alpha_i = b_i p$ and $\beta_i = d_i q$. In fact, using Theorem 1.19, we prove the following more general result, which is Theorem 4.4.

Theorem 1.30. *Assume A^0 and A^1 are nonproportional matrices of the form*

$$A^i := \begin{pmatrix} a_i & b_i \\ c_i & -a_i \end{pmatrix},$$

for $i = 0, 1$, where a_i, b_i, c_i are real numbers satisfying

$$a_i^2 + b_i c_i < 0.$$

Then all the average growth rates are equal and strictly positive:

$$\Lambda^+ = \Lambda^- > 0.$$

Thus, as an immediate corollary of Theorems 1.29 and 1.30, we deduce that the conjecture of Takeuchi et al. is true.

Corollary 1.4. *Assume that the vector fields F^0 and F^1 given by* (1.15) *are non collinear. Then for any $(x_0, y_0) \in \mathbb{R}^2_{++}$, with probability* 1,

$$\limsup x_t = \limsup y_t = +\infty, \quad \liminf x_t = \liminf y_t = 0.$$

1.6.3 The particular case where the zero is on a common invariant face

The results of this section come from my prepint [Str18], which has been submitted:

E. Strickler, *Randomly Switched Vector Fields sharing a Zero on a common invariant Face*, arXiv preprint : 1810.06331 (2018).

These results, which can be found in Chapter 5, are motivated by the following problem. Consider random switching between two Lorenz vector fields F^i, $i = 0, 1$:

$$F^i(x, y, z) = \begin{pmatrix} \sigma_i(y - x) \\ r_i x - y - xz \\ xy - b_i z \end{pmatrix}, \tag{1.18}$$

with $\sigma_0 = \sigma_1 = 10$, $b_0 = b_1 = 8/3$, $r_0 = 28$, and $r_1 \neq r_0$ close to 28. Note that 0 is a common equilibrium of F^0 and F^1. It has been known since the proof of Tucker [Tuc99] that F^0 admits a robust strange attractor Γ_0. Thus for r_1 close to r_0, F^1 shares this property. In [BH12], Bakhtin and Hurth showed that the compact set

$$M = \{(x,y,z) \in \mathbb{R}^3 \, : \, 2r_0\sigma(x^2+y^2) + 2\sigma b(z_0 - r_0)^2 \leq 2\sigma b r_0^2\}$$

is forward invariant, and that the strange attractor Γ_0 is accessible from every point that does not lie on the z-axis. Moreover they proved that the strong bracket condition holds at any point which is not on the z-axis. Then they argue that by compactness of M, there exists an invariant probability measure, and that it has to be absolutely continuous with respect to the Lebesgue measure due to the bracket condition. However, this argument is not sufficient : there exists indeed an invariant probability measure on M, which is $\delta_0 \otimes p$, where p is the invariant probability measure of the Markov chain $(I_t)_{t\geq 0}$ on $E = \{0,1\}$. However, this measure is not absolutely continuous. The main fallacy in [BH12] was that the authors didn't pay attention to the fact that M without the z - axis is no more compact. A natural idea to prove the existence of an absolutely continuous invariant probability measure would be to apply Theorem 1.25 above. Nonetheless, it does not apply because the F^i's have exactly two Lyapunov exponents, of opposite signs. Indeed, the Jacobian matrix of F^i at 0 has the block diagonal form :

$$A^i = \begin{pmatrix} B^i & 0 \\ 0 & -b_i \end{pmatrix},$$

where

$$B^i = \begin{pmatrix} -\sigma_i & \sigma_i \\ r_i & -1 \end{pmatrix}.$$

In particular, it is easily seen that $\nu = \delta_{(0,0,1)} \otimes p$ is an invariant probability measure of (Θ, J) on $S^2 \times E$, and that

$$\Lambda(\nu) = -(p_0 b_0 + p_1 b_1) < 0.$$

On the other hand, it can be proven that there exists another ergodic measure μ, supported on $\{\theta = (\theta_1, \theta_2, \theta_3) \in S^2 \, : \, \theta_3 = 0\} \times E$ and such that $\Lambda(\mu) > 0$. Thus none of the above theorems apply. However, note that the z - axis is invariant, and that the negative Lyapunov exponent is carried by this axis. This gives an intuition on why the persistence theorems given above should still hold.

Let us go back to the general case and be more precise. In previous setting of switched vector fields vanishing at 0, write $\mathbb{R}^d$ as $\mathbb{R}^n \times \mathbb{R}^m$ and assume that the face $\{0\} \times \mathbb{R}^m$ is invariant for every vector field F^i. That is, if we write $x \in \mathbb{R}^d$ as $x = (x_n, x_m)$ and $F^i(x) = (F^i_n(x), F^i_m(x))$, then for all $x_m \in \mathbb{R}^m$, $F^i_n(0, x_m) = 0$. This implies in particular that the face $\{0\} \times \mathbb{R}^m$ is invariant for the PDMP and that A^i is block lower triangular :

$$A^i = \begin{pmatrix} B^i & 0 \\ C^i & D^i \end{pmatrix}, \tag{1.19}$$

with $B^i \in M_n(\mathbb{R})$, $C^i \in M_{m,n}(\mathbb{R})$ and $D^i \in M_m(\mathbb{R})$. Now let $\Lambda(A)$, $\Lambda(B)$, and $\Lambda(D)$ denote the sets of average growth rates associated to the families $(A^i)_{i\in E}$, $(B^i)_{i\in E}$ and $(D^i)_{i\in E}$, respectively. We also let Λ_A^+ and Λ_A^- be the extremal values of these growth rates, and use similar notations for B and D. Then we have the following proposition, which summarised Lemma 5.1 and Proposition 5.2 :

Proposition 1.19. *With the above notation,* *1.* $\Lambda_A^+ = \max(\Lambda_B^+, \Lambda_D^+)$*;*

2. $\Lambda(D) \subset \Lambda(A)$*. In particular,* $\Lambda_A^- \leq \Lambda_D^-$*.*

Remark 1.16. *The first point of the above proposition is an immediate consequence of Theorem 1.18 and Proposition 1.18.*

From this proposition and Theorem 1.23, one deduce that if both Λ_B^+ and Λ_D^+ are negative, then so is Λ_A^+ and, provided 0 is accessible, one has for all $(x,i) \in M^* \times E$

$$\mathbb{P}_{(x,i)} \left(\limsup_{t \to \infty} \frac{1}{t} \|X_t\| \leq \Lambda_A^+ \right) = 1.$$

The next theorem is Theorem 5.2. It describes what happens when the Lyapunov exponents are of opposite signs. We set $M_0 = \{(x_n, x_m) \in M \,:\, x_n = 0\}$, $M_+ = M \setminus M_0$ and for $\delta > 0$, $M_0^\delta = \{(x_n, x_m) \in M_+ \,:\, \|x_n\| < \delta\}$.

Theorem 1.31. *Assume* $\Lambda_B^- > 0 > \Lambda_D^+$ *and that* 0 *is accessible from* $M_0 \times E$*. Then :*

(i) *For all* $\varepsilon > 0$*, there exists* $\delta > 0$ *such that for all* $x \in M_+$*,* $i \in E$*,* $\mathbb{P}_{x,i}$ *almost surely,*

$$\limsup_{t \to \infty} \Pi_t(M_0^\delta \times E) \leq \varepsilon.$$

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

(ii) *There exist positive constants* θ, K *such that for all* $\mu \in \mathcal{P}_{inv}(M_+ \times E)$*,*

$$\sum_{i \in E} \int \|x_n\|^{-\theta} d\mu^i(x_n, x_m) \leq K.$$

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

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

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

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

As in the previous section, the additional assumption that some bracket conditions hold leads to stronger statements, given in detail in Chapter 5. In particular, these theorems imply that the switching between Lorenz vector fields considered at the beginning of this section does indeed admit an invariant probability measure which is absolutely continuous. Furthermore, the law of the process converges exponentially fast in total variation to this invariant distribution.

1.6.4 A user-friendly condition for exponential ergodicity of randomly switched vector fields

Based on a joint work with Michel Benaïm and Tobias Hurth [BHS18].

M. Benaïm, T. Hurth and E. Strickler, *A user-friendly condition for exponential ergodicity in randomly switched environments*, Electronic Communications in Probability. **23** (2018), Paper No. 44, 12.

As explained in Section 1.2.4, one way to obtain ergodicity of a PDMP is to find an accessible point where the strong bracket condition holds. Moreover, in full generality, the weak bracket condition is not enough to ensure that the process converges to its unique stationary distribution.

From the definition, it is clear that it is more involving to check the strong bracket condition than the weak one. For example, when the number of vector fields is equal to the dimension of the space, it could be sufficient to check if the vector fields at a point constitute a free system to know that the weak bracket condition holds at that point; whereas for the strong bracket condition, one needs to compute at least the first-order Lie brackets.

In a recent paper [LLC17], Li, Liu and Cui showed that the existence of a globally asymptotically stable equilibrium for one of the vector fields, combined with the existence of an accessible point at which the weak bracket condition holds, is sufficient to have the convergence of the process in total variation to its unique stationary distribution.

Together with Michel Benaïm and Tobias Hurth, we proved that the assumption on the equilibrium can be widely relaxed : it is sufficient that there exists an accessible point at which a barycentric combination of the vector fields vanishes. Moreover, using the results in [BLBMZ15] quoted in Section 1.2.4, we simplified the proof in [LLC17] and showed that the convergence holds at an exponential rate. Namely :

Theorem 1.32. *Suppose that*

(i) *There exist $\alpha_1, \dots, \alpha_N \in \mathbb{R}$ with $\sum \alpha_i = 1$ and $e^\star \in M$ such that $\sum \alpha_i F^i(e^\star) = 0$,*

(ii) *There exists a point x^* accessible from $\{e^\star\}$ where the weak bracket condition holds.*

Then for all $j \in E$, $(e^\star, j)$ is a Doeblin point.

In view of the general results in Section 1.1, we obtain immediately the next corollary:

Corollary 1.5. *In addition to the assumptions in Theorem 1.32, suppose that $e^\star$ is accessible. and that M is compact. Then, the process Z admits a unique invariant probability measure π which is absolutely continuous with respect to Lebesgue measure. Moreover, there exist positive constants C, γ such that for all $t \geq 0$ and for all $(x,i) \in M \times E$,*

$$\|P_t((x,i),\cdot) - \pi\|_{TV} \leq C e^{-\gamma t}.$$

As an application of Theorem 1.32, we prove that point *4.(d)* of Theorem 1.15 holds even though $\frac{\beta_0 \alpha_1}{\alpha_0 \beta_1} = \frac{a_0 c_1}{c_0 a_1} = \frac{b_0 d_1}{d_0 b_1}$ (see Remark 1.10).

It is natural to wonder if conditions **(i)** and **(ii)** imply the strong bracket condition. In Chapter 2, we show that it is not the case in general, but that it is true for analytic vector fields. We also exhibit an example of vector fields satisfying the strong bracket condition but not condition **(i)**.

1.6.5 Approximation with extinction of Markov Processes that never die

This section presents the results obtained in Chapter 6. It contains only partial results, and has not been submitted on a archive platform. All the processes with an extinction set described above share the same feature : *they never die in finite time.* That is, whenever $X_0 \neq 0$, then $X_t \neq 0$ for all $t > 0$. A natural question is whether this assumption is realistic in ecology, population models etc. At a first glance, it is not : death in finite time is the tragic destiny of every population. However, as the size of the population gets very large, the time of extinction can be very large, and we might not be too far from real life by assuming that the population never gets extinct. In Chapter 6, we try to legitimate this intuition. Given an immortal Markov process $(X_t)_{t\geq 0}$ living on a compact metric space M, we consider a family of Markov processes $(X_t^N)_{t\geq 0}$ that die in finite time and approximate X. More precisely, we assume that :

Assumption 1.7. *There exists a closed set $M_0 \subset M$ such that*

(i) *M_0 is invariant for X : for all $t \geq 0$, $X_t \in M_0 \Leftrightarrow X_0 \in M_0$,*

(ii) *M_0 is absorbing for X^N : for all $t, s \geq 0$, $X_t^N \in M_0 \Rightarrow X_{t+s}^N \in M_0$;*

(iii) *For all $x \in M$, $\mathbb{P}_x(T_0^N < \infty) = 1$, where $T_0^N = \inf\{t \geq 0 \,:\, X_t^N \in M_0\}$;*

(iv) *For all $x \in M$ and $t > 0$, $\mathbb{P}_x(T_0^N > t) > 0$.*

(v) *For all N, the process X^N admits a QSD α^N.*

(vi) *For all $T, \delta > 0$,*

$$\lim_{N\to\infty} \max_{x\in M} \mathbb{P}_x \left(\sup_{t\in[0,T]} d(X_t^N, X_t) > \delta \,\middle|\, X_0^N = X_0 = x \right) = 0. \tag{1.20}$$

The process X belongs to the class of processes introduced in Section 1.3, whereas X^N belongs to the one introduced in Section 1.5. The convergence in (1.20) is quite strong. It appears for example in [Kur81], when X^N is a Markov chain on a finite grid of size N and X solves an ODE.

For $N \geq 0$, let λ_N be the extinction rate associated to the QSD α^N (see Proposition 1.16 above). The questions that we want to answer are the following. Can we say something about the behaviour of λ_N as N goes to infinity ? What can be said about the weak limit points of the sequence $(\alpha^N)_{N\geq 0}$? As one could expect, it turns out that the answers to these questions depend on the behaviour of X near the extinction set. In particular, we are able to say something if the process is H - persistent, as defined in Section 1.3. The following theorem sums up the partial results obtained in Chapter 6.

Theorem 1.33. *Let α denote a weak limit point of α^N. Then :*

1. *If X is H-persistent, $\lim_{N\to\infty} \lambda_N = 0$ and α is an invariant probability measure of X;*

2. *If X is H-nonpersistent and M_0 is accessible from M, then $\alpha(M_0) = 1$.*

The first part of the above theorem legitimates the use of a process X which never dies in finite time to approximate real-life situation : if N is very large, λ_N is small and thus the mean time to extinction is huge. Nonetheless, the main remaining question is whether α puts all its mass on M_+ or if in the limit, some mass could escape to M_0.

We give an application to the spread of a disease in a random environment, whose limiting process is the SIS PDMP from Section 1.6.1. That is, for $N \geq 1$, we consider a Markov chain (X^N, I^N) on the space $M_N = \{0, \frac{1}{N_1}, \dots, 1\} \times \dots \times \{0, \frac{1}{N_d}, \dots, 1\} \times E$,. Here N represents the total population, and N_i the population into group i. The transition rates are given for all $(x,k) \in M_N$ by

$$\begin{aligned}(x,k) \longrightarrow &(x + \frac{e_i}{N_i}, k) \quad \text{at rate} \quad N\frac{N_i}{N}(1-x_i)\sum_{j=1}^{d} C_{ij}^k x_j \\ &(x - \frac{e_i}{N_i}, k) \quad \text{at rate} \quad N\frac{N_i}{N} - D_i^k x_i \\ &(x, l) \qquad\quad \text{at rate} \quad a_{k,l}(x).\end{aligned}$$

Here $(e_1, \dots, e_d)$ stands for the canonical basis of $\mathbb{R}^d$, and x_i is the proportion of infected individual in group i. We assume that for all i, the proportion $\frac{N_i}{N}$ converges to some $p_i \in (0,1)$. We show in Chapter 6 that (X^N, I^N) and the PDMP (X, I) with infinitesimal generator

$$Lf(x,k) = \langle F^k(x), \nabla f(x,k) \rangle + \sum_{l \in E} a_{k,l}(x) \left[f(x,l) - f(x,k) \right]$$

satisfy Assumption 1.7. Here F^k is the Lajmonovich and Yorke vector field given by

$$F_i^k(x) = (1 - x_i) \sum_{j=1}^{d} C_{ij}^k x_j - D_i^k x_i.$$

A second application is given to a killed process, where the killing rate goes to 0.

1.7 Structure of the thesis

This thesis contains six chapters, including the present introduction. Chapter 2 reproduces the core of the article [BHS18] published at **Electronic Communication in Probability**, Chapter 3 reproduces [BS19], accepted for publication at ***Annals of Applied Probability***. Chapter 4 contains the core of the article [HS19], accepted at **SIAM Journal on Mathematical Analysis**. Chapter 5 has led to the article [Str18]. Chapter 6 is concerned with the results that I have obtained so far in the situation of approximation with extinction of Markov processes that never die.

Chapter 2

A user-friendly condition for exponential ergodicity in randomly switched environments

In this Chapter, we reproduce the article [BHS18], that has been published in *Electronic Communication in Probability*. Together with Michel Benaïm and Tobias Hurth, we give a slightly different condition for the exponential convergence of a PDMP towards its invariant probability (see Corollary 1.2). More precisely, we show that if in addition to the weak bracket condition, there exists an accessible point at which a barycentric combination of the vector fields vanishes, the convergence holds. We also prove that this condition implies the strong bracket condition if the vector fields are analytic, and give a counterexample in the non-analytic case.

This result was inspired by a work by Li, Liu and Cui [LLC17].

Keywords: Piecewise deterministic Markov processes; random switching; Hörmander-bracket conditions; ergodicity; stochastic persistence

MSC primary: 60K35, 60G17, 60J60

Contents

2.1 Introduction

Let $E = \{1, \dots, N\}$ be a finite set and $\mathrm{F} = \{F^i\}_{i \in E}$ a family of smooth globally integrable vector fields on $\mathbb{R}^d$. For each $i \in E$ we let $\varphi^i = \{\varphi^i_t\}$ denote the flow induced by F^i. We assume throughout that there exists a compact set $M \subset \mathbb{R}^d$ which is *positively invariant* under each φ^i. That is

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

for all $t \geq 0$. Our assumption that $M \subset \mathbb{R}^d$ is mostly for convenience. The results of this chapter can readily be generalized to the situation where M is a subset of a finite-dimensional smooth manifold.

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

$$\mathcal{L}g(x,i) = \langle F^i(x), \nabla g^i(x) \rangle + \sum_{j \in E} a_{ij}(x)(g^j(x) - g^i(x)), \tag{2.1}$$

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 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

$$\frac{dX_t}{dt} = F^{I_t}(X_t), \tag{2.2}$$

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

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

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

The process Z belongs to the class of processes called Piecewise Deterministic Markov Processes (PDMP), introduced by Davis in [Dav84]. Ergodic properties of these processes have recently been the focus of much attention (e.g., [BLBMZ12], [CH15],[BLBMZ15], [BCL17], [BS19], [BHLM18]).

Recall that if there exists an accessible point at which the *weak bracket condition* holds (cf. Definitions 1.15 and 1.16) , the process admits a unique invariant probability measure which is absolutely continuous with respect to the Lebesgue measure on $M \times E$ (see Proposition 1.13). If the weak bracket condition is replaced by the *strong bracket condition* (cf. Definition 1.16), the process then converges in total variation (see Corollary 1.2). Simple examples show that the weak bracket condition itself is not sufficient to ensure convergence (cf. [BH12]).

Recently, Li, Liu and Cui showed in [LLC17] that the two following conditions yield convergence in total variation (see [LLC17, Theorem 9]) :

(i') There exists a globally asymptotically stable (G.A.S.) equilibrium for one of the vector fields,

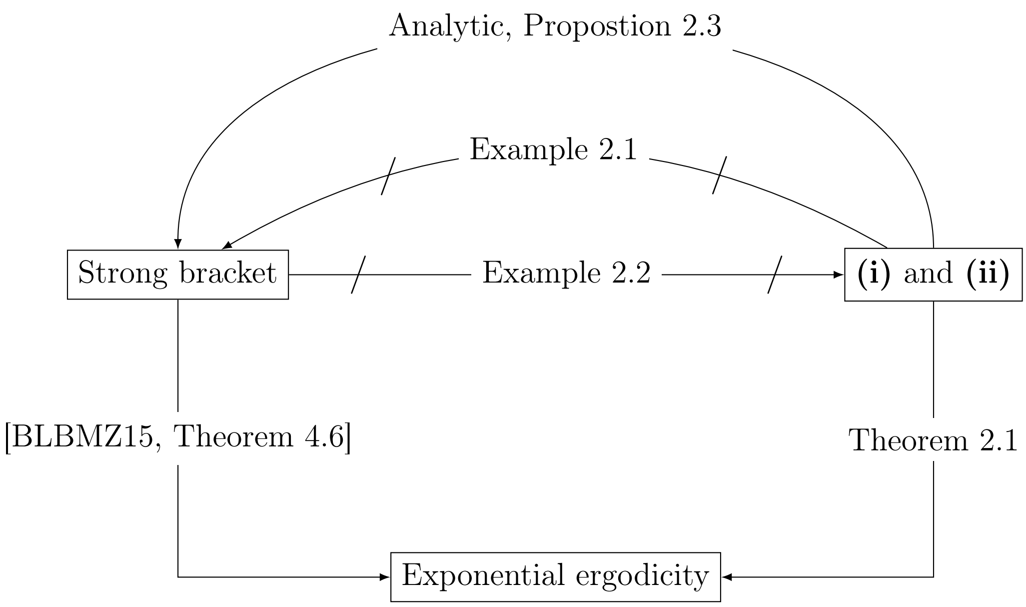
(ii) The weak bracket condition holds at an accessible point.

In this chapter, we replace **(i')** by the more general condition

(i) There exists an accessible point $e^\star$ at which a barycentric combination of the vector fields vanishes,

and prove exponential convergence in total variation (see Theorem 2.1 and Corollary 2.1). Our proof is inspired by [LLC17] but is simplified using results of [BLBMZ15].

It turns out that when the vector fields are analytic, **(i)** and **(ii)** imply the strong bracket condition at $e^\star$ (cf. Proposition 2.3). Nonetheless, **(i)** and **(ii)** are usually much easier to verify than the strong bracket condition. This is illustrated by the examples in Section 2.3. In the nonanalytic case, neither condition implies the other as shown in Section 2.2.2 (see Examples 2.1 and 2.2). All these results are summarized in the following scheme.



2.2 Definitions and main results

For convenience, we recall the following definitions. The reader is referred to Chapter 1 for more details.

For $\mathbf{i} = (i_1, \ldots, i_m) \in E^m$ and $\mathbf{u} = (u_1, \ldots, u_m) \in \mathbb{R}_+^m$, we denote by $\mathbf{\Phi}_{\mathbf{u}}^{\mathbf{i}}$ the composite flow : $\mathbf{\Phi}_{\mathbf{u}}^{\mathbf{i}} = \varphi_{u_m}^{i_m} \circ \ldots \circ \varphi_{u_1}^{i_1}$. For $x \in M$ and $t \geq 0$, we denote by $\gamma_t^+(x)$ (resp. $\gamma^+(x)$) the set of points that are reachable from x at time t (resp. at any nonnegative time) with a composite flow:

$$\gamma_t^+(x) = \{\mathbf{\Phi}_{\mathbf{v}}^{\mathbf{i}}(x),\ (\mathbf{i}, \mathbf{v}) \in E^m \times \mathbb{R}_+^m, m \in \mathbb{N}, v_1 + \ldots + v_m = t\},$$

$$\gamma^+(x) = \bigcup_{t \geq 0} \gamma_t^+(x).$$

Definition 2.1. *A point $x^* \in M$ is* $\{F^i\}$-accessible *from $B \subset M$ if $x^* \in \cap_{x \in B} \overline{\gamma^+(x)}$.*

From now on, we let $(P_t)_{t\geq 0}$ be the semigroup induced by $(Z_t)_{t\geq 0}$ on $\mathcal{M} = M \times E$. Because of the irreducibility assumption on the rate matrix $a(x)$, Definitions 2.1 coïncide with the general definition of accessibility given for Markov processes in Section 1.1 (see Proposition 1.10). Therefore, in the sequel, we will say that a point $x^* \in M$ is accessible from $B \subset M$ if it is $\{F^i\}$-accessible from B. We will simply say that x^* is accessible if it is $\{F^i\}$-accessible from M. Set $\mathrm{F}_0 = \{F^i\}_{i\in E}$, $\mathrm{F}_{k+1} = \mathrm{F}_k \cup \{[F^i, V], V \in \mathrm{F}_k\}$, $\mathcal{F}_0 = \{F^i - F^j \ : i, j = 1, \dots m\}$ and $\mathcal{F}_{k+1} = \mathcal{F}_k \cup \{[F^i, V] \ : V \in \mathcal{F}_k\}$. Here $[\cdot,\cdot]$ stands for the Lie bracket operation, which is defined as

$$[V, W](x) = DW(x)V(x) - DV(x)W(x), \quad x \in \mathbb{R}^d,$$

for smooth vector fields V and W on $\mathbb{R}^d$ with differentials DV and DW. The following definition is given in [BLBMZ15] (see Definition 1.16 in Chapter 1).

Definition 2.2. *We say that the* weak bracket *(resp.* strong bracket*) condition holds at $p \in M$ if the vector space spanned by the vectors $\{V(p) \ : V \in \cup_{k\geq 0}\mathrm{F}_k\}$ (resp. $\{V(p) \ : V \in \cup_{k\geq 0}\mathcal{F}_k\}$) has full rank.*

It is clear from this definition that the strong bracket condition implies the weak one. Weak bracket and strong bracket conditions are equivalent to Condition B and Condition A in [BH12], respectively. The weak bracket condition is closely related to the classical Hörmander hypoellipticity condition that yields smoothness of transition densities for diffusions (see e.g. [Nua06]). More background on the weak and strong bracket conditions with an emphasis on how they relate to controllability is provided in [SJ72].

2.2.1 Main result

We now state our main result.

Theorem 2.1. *Suppose that*

(i) *There exist $\alpha_1, \dots, \alpha_N \in \mathbb{R}$ with $\sum \alpha_i = 1$ and $e^\star \in M$ such that $\sum \alpha_i F^i(e^\star) = 0$,*

(ii) *There exists a point x^* accessible from $\{e^\star\}$ where the weak bracket condition holds.*

Then for all $j \in E$, $(e^\star, j)$ is a Doeblin point.

Note that we do not impose that the α_i are nonnegative. In particular, condition **(i)** holds whenever two vector fields at some point are collinear but not equal.

The following corollary is a consequence of standard results (see Corollary 1.1 and Remark 1.4).

Corollary 2.1. *In addition to the assumptions in Theorem 2.1, suppose that $e^\star$ is accessible. Then, the process Z admits a unique invariant probability measure π which is absolutely continuous with respect to Lebesgue measure. Moreover, there exist positive constants C, γ such that for all $t \geq 0$ and for all $(x, i) \in M \times E$,*

$$\|P_t((x, i), \cdot) - \pi\|_{TV} \leq Ce^{-\gamma t}.$$

In Section 2.3, we give more applications in a stochastic persistence context, relying on recent results in [Ben18] (see Section 1.3). Theorem 2.1 is a direct consequence of Theorem 1.11 and of Proposition 2.1 that we state below. Here and throughout, for $s > 0$ and $m \in \mathbb{N}^*$, we set $D^s_m = \{\mathbf{v} \in \mathbb{R}^m_+ \ : v_1 + \ldots + v_m \leq s\}$.

Proposition 2.1. *Under conditions **(i)** and **(ii)** of Theorem 2.1, there exist $s > 0$, $i_{m+1} \in E$, $\mathbf{i} \in E^m$ and $\mathbf{u} \in \mathbb{R}^m_+$ with $u_1 + \ldots + u_m < s$ such that the map $\Psi : D^s_m \to \mathbb{R}^d$, $\mathbf{v} \to \varphi^{i_{m+1}}_{s-(v_1+\ldots+v_m)} \circ \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(e^\star)$ is a submersion at $\mathbf{u}$.*

2.2.2 Links with the strong bracket condition

In [BLBMZ15] and [BH12], the authors show that the conclusions of Theorem 2.1 and Corollary 2.1 hold when the weak bracket condition is replaced by the strong one (see Corollary 1.2). A natural question is whether our assumptions already imply that the strong bracket condition holds at some point. We address this question in Propositions 2.2 and 2.3.

Proposition 2.2. *Let $e^\star \in M$ satisfy condition* **(i)** *of Theorem 2.1. Suppose further that the weak bracket condition holds at $e^\star$. Then, the strong bracket condition is also satisfied at $e^\star$.*

Proof To simplify notation, we set

$$W(e^\star) = \{V(e^\star) : V \in \cup_{k\geq 0}\mathrm{F}_k\}, \quad S(e^\star) = \{V(e^\star) : V \in \cup_{k\geq 0}\mathcal{F}_k\}.$$

We will show that the linear spans of $W(e^\star)$ and $S(e^\star)$ are equal to each other, which then implies the proposition. It is clear that the span of $S(e^\star)$ is a subspace of the span of $W(e^\star)$. Therefore, it suffices to show that $W(e^\star)$ is contained in the span of $S(e^\star)$. Fix a vector field $V \in \cup_{k\geq 0}\mathrm{F}_k$ and let j be the smallest nonnegative integer such that $V \in \mathrm{F}_j$. By induction it is not hard to see that for any $i \geq 1$, the collection of vector fields $\mathrm{F}_i \setminus \mathrm{F}_{i-1}$ is contained in the span of $\cup_{k\geq 0}\mathcal{F}_k$. Thus, if $j \geq 1$, the point $V(e^\star)$ lies in the span of $S(e^\star)$. If $j = 0$, there is $l \in E$ such that $V = F^l$. By condition **(i)**, there are real numbers $(\alpha_i)_{i\in E}$ such that $\sum_{i\in E} \alpha_i = 1$ and $\sum_{i\in E} \alpha_i F^i(e^\star) = 0$. Therefore,

$$F^l(e^\star) = \sum_{i\in E} \alpha_i F^l(e^\star) - \sum_{i\in E} \alpha_i F^i(e^\star) = \sum_{i\in E} \alpha_i (F^l(e^\star) - F^i(e^\star)).$$

Since the vector fields $(F^l - F^i)_{i\in E}$ lie in $\mathcal{F}_0$, we have again that $V(e^\star)$ is in the span of $S(e^\star)$. This finishes the proof. □

Proposition 2.3. *Assume that for all $i \in E$, F^i is analytic and that the assumptions of Theorem 2.1 hold. Then $e^\star$ satisfies the strong bracket condition.*

In most applications, the vector fields governing the PDMP are analytic (see also Section 2.3). As a consequence, the interest of Theorem 2.1 lies essentially in the fact that the weak bracket condition is easier to verify than the strong one. The proof of Proposition 2.3 relies on the following result, due to Sussmann and Jurdjevic [SJ72, Corollary 4.7].

Theorem 2.2 (Sussmann – Jurdjevic)**.** *Assume that the vector fields $(F^i)_{i\in E}$ are analytic, and let x be any point in M. Then, there is $t > 0$ such that $\gamma_t^+(x)$ has nonempty interior if and only if the strong bracket condition holds at x.*

Proof of Proposition 2.3
By Proposition 2.1, there are $s > 0$, $i_{m+1} \in E$, $\mathbf{i} \in E^m$ and $\mathbf{u} \in \mathbb{R}_+^m$ with $u_1 + \ldots + u_m < s$ such that $\Psi : \mathbf{v} \to \varphi^{i_{m+1}}_{s-(v_1+\ldots+v_m)} \circ \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(e^\star)$ is a submersion at $\mathbf{u}$. By the constant-rank theorem, there exists an open neighborhood U of $\mathbf{u}$ such that $\Psi(U)$ is open. Without loss of generality, we can assume that $v_1 + \ldots + v_m < s$ for all $\mathbf{v} \in U$. Then, $\Psi(U)$ is a nonempty open subset of $\gamma_s^+(e^\star)$. By Theorem 2.2, $e^\star$ satisfies the strong bracket

condition. □

From a more theoretical point of view, we now provide an example in the plane where conditions **(i)** and **(ii)** are satisfied, but, in the absence of analyticity, there is no point where the strong bracket condition holds.

Example 2.1. We work in polar coordinates (θ, r). On the annulus

$$M = \left\{(\theta, r) : \tfrac{1}{2} \leq r \leq 2\right\},$$

we switch between vector fields $F^0(\theta, r) = (1, h(r))^T$ and $F^1(\theta, r) = (f(\theta), g(\theta) + h(r))^T$, where

$$h(r) = r(1 - r),$$

and where f and g satisfy the following properties:

1. The functions f and g are C^∞ and 2π-periodic on $\mathbb{R}$.
2. We have $0 < f \leq 1$ and $0 \leq g \leq 1$.
3. We have $f(\frac{\pi}{2}) = \frac{1}{2}$ and $g(0) > 0$. Moreover, there is $\epsilon \in (0, \frac{\pi}{4})$ such that $f(\theta) = 1$ for $|\theta - \frac{\pi}{2}| > \epsilon$ and $g(\theta) = 0$ for $|\theta| > \epsilon$.

It is easy to see that such functions f and g exist and that they cannot be analytic. Also note that M is positively invariant under the flows associated with F^0 and F^1 because $h(\frac{1}{2}) > 0$ and $g(\theta) + h(2) < 0$ for all θ. Since M is compact and since f, g and h are smooth functions, the vector fields F^0 and F^1 are globally integrable.

The point $e^\star = (\frac{\pi}{2}, 1)^T$ is an equilibrium point of the vector field $2F^1 - F^0$, so condition **(i)** is satisfied. Since $h(r) > 0$ for $r \in (0, 1)$ and $h(r) < 0$ for $r > 1$, the unit circle is a global attractor of F^0. Thus, any point on the unit circle, in particular the point $e^\star$, is accessible from any starting point in M. The weak bracket condition holds at the point $(0, 1)^T$ because $F^0(0, 1) = (1, 0)^T$ and $F^1(0, 1) = (1, g(0))^T$ generate the entire tangent space at $(0, 1)^T$. As $(0, 1)^T$ lies on the unit circle, it is accessible from $e^\star$.

It remains to show that the strong bracket condition is nowhere satisfied. We have

$$[F^0, F^1](\theta, r) = (f'(\theta), g'(\theta) - h'(r)g(\theta))^T$$

and

$$F^1(\theta, r) - F^0(\theta, r) = (f(\theta) - 1, g(\theta))^T.$$

If $|\theta - \frac{\pi}{2}| > \epsilon$, both $[F^0, F^1](\theta, r)$ and $(F^1 - F^0)(\theta, r)$ have θ-coordinate 0. And if $|\theta| > \epsilon$, the r-coordinate of $[F^0, F^1]$ and $F^1 - F^0$ vanishes. Now, let $k(\theta, r)$ be a smooth function and let $K_i(\theta, r) = k(\theta, r)(1 - i, i)^T$ for $i \in \{0, 1\}$. Then,

$$\begin{aligned}
[F^0, K_1](\theta, r) &= (0, *)^T, & [F^0, K_0](\theta, r) &= (*, 0)^T, \\
[F^1, K_1](\theta, r) &= (0, *)^T, & [F^1, K_0](\theta, r) &= (*, -g'(\theta)k(\theta, r))^T,
\end{aligned}$$

and $g'(\theta)k(\theta, r) = 0$ for $|\theta| > \epsilon$. Here, $*$ stands for some term, possibly depending on θ and r, that may differ from equation to equation. This shows that for any $(\theta, r) \in M$, $V(\theta, r)$ lies in the linear span of $(1, 0)^T$ for all $V \in \cup_{k \geq 0} \mathcal{F}_k$, or $V(\theta, r)$ lies in the linear span of $(0, 1)^T$ for all $V \in \cup_{k \geq 0} \mathcal{F}_k$. It follows that the strong bracket condition doesn't hold at any point $(\theta, r) \in M$.

In the previous example, the origin had to be excluded from M in order to ensure that the unit circle is globally accessible. It could be interesting to determine whether there are PDMPs for which conditions **(i)** and **(ii)** are satisfied, the strong bracket condition nowhere holds, and M is simply connected.

As illustrated by the following example, the strong bracket condition does not imply condition **(i)**, not even if the vector fields are analytic.

Example 2.2. On the two-dimensional torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, we switch between $F^0(x, y) = (1, 0)^T$ and $F^1(x, y) = (0, 1 + \epsilon \sin(2\pi x))^T$, where $\epsilon > 0$ is small. Any point in $\mathbb{T}^2$ can then be reached from any starting point. For $\alpha \in \mathbb{R}$, we have

$$\alpha F^0(x, y) + (1 - \alpha)F^1(x, y) = (\alpha, (1 - \alpha)(1 + \epsilon \sin(2\pi x)))^T,$$

which is never zero. However,

$$[F^0, F^1](x, y) = (0, -\epsilon 2\pi \cos(2\pi x))^T,$$

so the vectors $[F^0, F^1](0, 0)$ and $F^0(0, 0) - F^1(0, 0) = (1, -1)^T$ span the tangent space at $(0, 0)$, and the strong bracket condition is satisfied.

2.3 Applications

In this section, we give some applications of Theorem 2.1 in the context of population models with an extinction set. For a general framework on Markov models with an extinction set, the reader is referred to Section 1.3 and [Ben18].

The following theorem is an immediate consequence of Theorems 1.13 and 2.1.

Theorem 2.3. *Assume that conditions* **(i)** *and* **(ii)** *hold, that Z is H - persistent and that $e^\star$ is accessible from M_+. Then Z admits a unique invariant probability measure Π on $M_+ \times E$ and there exist $\theta, C, \gamma > 0$ such that for all $t \geq 0$ and for all $(x, i) \in M_+ \times E$,*

$$\|P_t((x, i), \cdot) - \Pi\|_{TV} \leq C \left(1 + e^{\theta V(x,i)}\right) e^{-\gamma t}.$$

An application is given in Chapter 3 for SIS model in random environment and in the following subsection.

2.3.1 Lotka-Volterra in random environment

In this section, we consider the competitive Lotka-Volterra model in a fluctuating environment studied in [BL16] and described in Section 1.3.2, and show how our method can be used to improve one of their results. More precisely, for $i \in \{0, 1\}$, let F^i be defined as

$$F^i(x, y) = \begin{pmatrix} \alpha_i x(1 - a_i x - b_i y) \\ \beta_i y(1 - c_i x - d_i y) \end{pmatrix}, \tag{2.3}$$

with $\alpha_i, \beta_i, a_i, b_i, c_i, d_i > 0$. For $\eta > 0$ small enough, the flows φ_t^i leave positively invariant the compact set $K = \{(x, y) \in \mathbb{R}_+^2 \ : \eta \leq x + y \leq 1/\eta\}$, and the extinction set K_0 is the union of $K_0^{\mathbf{x}} = \{(x, y) \in M \ : x = 0\}$ and $K_0^{\mathbf{y}} = \{(x, y) \in M \ : y = 0\}$. It is shown in [BL16] (see also Theorem 1.15) that the long-term behavior of the process $(Z_t)_{t\geq 0} = (X_t, Y_t, I_t)_{t\geq 0}$ is determined by the sign of the invasion rates :

$$\Lambda_y = \int \beta_i(1 - c_i x)\mathrm{d}\mu(x, i),$$

and

$$\Lambda_x = \int \alpha_i(1 - b_i y)\mathrm{d}\hat{\mu}(y, i),$$

where μ and $\hat{\mu}$ are the unique invariant probability measures of the process Z restricted to $K_0^{\mathbf{y}} \times E$ and $K_0^{\mathbf{x}} \times E$, respectively. Recall from Section 1.3.2 that Z is H- persistent if and only if $\Lambda_x > 0$ and $\Lambda_y > 0$.

It is shown in [BL16] that if $\Lambda_x > 0$ and $\Lambda_y > 0$, then the process admits a unique invariant probability measure Π in $K_+ \times E$. But to show the convergence in total variation of the law of Z_t toward Π, the authors needed to check that the strong bracket condition is satisfied at some accessible point. They proved, except in the particular case where $\frac{\beta_0\alpha_1}{\alpha_0\beta_1} = \frac{a_0c_1}{c_0a_1} = \frac{b_0d_1}{d_0b_1}$, that this condition holds by using a formal calculus program. Thanks to Theorem 2.3, we withdraw this condition, and give an easier proof for the convergence in total variation.

In [BL16], of particular importance is the study of the averaged vector fields $F^s := sF^1 + (1-s)F^0$, for $s \in [0, 1]$. The vector field F^s is still a competitive Lotka-Volterra system of the form (2.3), with coefficients $\alpha_s, \beta_s, a_s, b_s, c_s, d_s$ that are barycentric combinations of the coefficients appearing in F^0 and F^1. The dynamics of the deterministic system generated by F^s depends on the position of s with respect to the two following (possibly empty) intervals:

$$I = \{s \in (0, 1) \ : a_s > c_s\}$$

and

$$J = \{s \in (0, 1) \ : b_s > d_s\}.$$

There are four regions of interest :

- $s \in (\overline{I})^c \cap (\overline{J})^c$: the equilibrium $(1/a_s, 0)$ is a global attractor for solutions with $x_0 \neq 0$;
- $s \in I \cap J$: the equilibrium $(0, 1/b_s)$ is a global attractor for solutions with $y_0 \neq 0$;
- $s \in I \cap (\overline{J})^c$: F^s admits a unique G.A.S. equilibrium $e_s \in M_+$;
- $s \in (\overline{I})^c \cap J$: F^s admits a unique equilibrium $e_s \in M_+$, which is a saddle whose stable manifold separates the basins of attraction of $(1/a_s, 0)$ and $(0, 1/b_s)$.

Here, $(\overline{I})^c$ and $(\overline{J})^c$ stand for the complement of the closure of I and J, respectively. The following proposition is a consequence of [BL16, Proposition 2.3 and Theorem 4.1].

Proposition 2.4. *Assume $\Lambda_y > 0$. Then $I \neq \emptyset$ and there exists a point m accessible from M_+ such that the weak bracket condition holds at m.*

From this proposition, we can derive the next lemma:

Lemma 2.1. *Assume $\Lambda_y > 0$. Then there exists $s \in [0, 1]$ such that F^s admits an equilibrium $e_s \in M_+$ which is accessible from M_+. In particular, condition* **(i)** *holds.*

This lemma combined with Proposition 2.4 and Theorem 2.3 implies the following corollary, which slightly improve [BL16, Theorem 4.1 - (iv)] (see Theorem 1.15).

Corollary 2.2. *Assume $\Lambda_y > 0$ and $\Lambda_x > 0$. Then there exist $C, \gamma, \theta > 0$ such that for all $t \geq 0$ and for all $(x, y, i) \in M_+ \times E$,*

$$\|P_t((x,i),\cdot) - \pi\|_{TV} \leq C\left(1 + \frac{1}{\|x\|^\theta} + \frac{1}{\|y\|^\theta}\right) e^{-\gamma t}.$$

Proof of Lemma 2.1 Since $\Lambda_y > 0$, I is nonempty by Proposition 2.4. Then we have three cases: either $I \cap J^c$ is nonempty, or I is a strict subset of J or $I = J$. We prove the lemma in these three cases. Assume first that $I \cap J^c \neq \emptyset$ and take $s \in I \cap J^c$. Then F^s admits a G.A.S. equilibrium $e_s \in M_+$, in particular it is accessible. Assume now that I is a strict subset of J. In particular, $I^c \cap J$ and $I \cap J$ are nonempty. Pick $s \in I^c \cap J$, then F^s admits a unique equilibrium $e_s \in M_+$, which is a saddle whose stable manifold W_s separates the basins of attraction of $(1/a_s, 0)$ and $(0, 1/b_s)$. We show that e_s is accessible. Choose a point $(x, y) \in M_+$. Then, if (x, y) is above W_s, follow the flow φ^0. As the resulting trajectory converges to $(1/a_0, 0)$, it needs to cross W_s. If (x, y) is below W_s, one can find a trajectory leading to $(0, 1/b_u)$ for some $u \in I \cap J$. In particular, this trajectory also crosses W_s. As e_s is also accessible from every point in W_s, it is accessible from everywhere in M_+. Finally, assume that $I = J = (s_1, s_2)$. Then the vector field F^{s_1} is of the form

$$F^{s_1}(x,y) = \begin{pmatrix} \alpha x(1 - ax - by) \\ \beta y(1 - ax - by) \end{pmatrix},$$

with $a = a_{s_1} = c_{s_1}$ and $b = b_{s_1} = d_{s_1}$. In particular, the line $y = 1/b(1 - ax)$ is composed of equilibria of F^{s_1}. Moreover, $(1/a_0, 0)$ and $(1/a_1, 0)$ lie on opposite sides of this line. Now we know by Proposition 2.4 that there exists an accessible point $m \in M_+$. Hence, depending on the position of m with respect to the line $y = 1/b(1 - ax)$, follow either φ^0 or φ^1 in order to cross the line when starting at m. Then the point where the line is crossed is accessible from m and therefore from M_+. □

2.4 Proof of Proposition 2.1

To prove Proposition 2.1, we will use Theorem 1.9. The following proposition is the key point of the proof :

Proposition 2.5. *Under the hypothesis of Theorem 2.1, there exist $s > 0$, $i \in E$, $\mathbf{i} = (i_1, \ldots, i_n) \in E^n$ and $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}_+^n$ with $s > u_1 + \ldots + u_n$ such that the map $\Psi : D_{n+1}^s \to \mathbb{R}^d$, $(\mathbf{v}, t) \to \varphi^i_{s - \sum v_i - t} \circ \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(e^\star)$ is a submersion at $(\mathbf{u}, 0)$.*

This proposition remains valid if we replace $e^\star$ by any point in M from which one can access a point x^* where the weak bracket condition holds. In particular, it is independent of our assumption that $e^\star$ is an equilibrium of a vector field of the form $\sum \alpha_i F^i$. The proposition is a consequence of the two lemmas we give now.

Lemma 2.2. *Suppose that there exists a point x^* accessible from $e^\star$ such that the weak bracket condition holds at x^*. Then there exists $(\bar{\mathbf{i}}, \bar{\mathbf{u}})$ such that the weak bracket condition holds at $\mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{u}}}(e^\star)$.*

Proof By Proposition 1.10, x^* is accessible from $e^\star$ if and only if $x^* \in \overline{\gamma^+(e^\star)}$. By continuity of the determinant and regularity of the vector fields, the weak bracket condition is an open condition. Thus if it holds at a point of $\overline{\gamma^+(e^\star)}$, it also holds at a point in $\gamma^+(e^\star)$, hence the result. □

Thanks to this lemma, we assume from now on that there exist $\bar{\mathbf{i}} = (\bar{i}_1, \ldots, \bar{i}_p)$ and $\bar{\mathbf{u}} = (\bar{u}_1, \ldots, \bar{u}_p)$ such that $x^* = \mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{u}}}(e^\star)$. Since x^* satisfies the weak bracket condition, Theorem 1.9 implies that there exists $m \geq d$, $\mathbf{i} = (i_1, \ldots, i_m) \in E^m$ and $\mathbf{u} = (u_1, \ldots, u_m) \in \mathbb{R}^m_+$ such that the map $\psi : \mathbf{v} \to \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x^*)$ is a submersion at $\mathbf{u}$. We denote $\mathbf{i}_- = (i_1, \ldots, i_{m-1})$ and $\mathbf{v}_- = (v_1, \ldots, v_{m-1})$, and for all $s > 0$, we define the map $\Psi^s : D^s_{m+p} \to \mathbb{R}^d$ by

$$\Psi^s : (\mathbf{v}_-, \bar{\mathbf{v}}, t) \to \varphi^{i_m}_{s-(v_1+\ldots+v_{m-1}+\bar{v}_1+\ldots+\bar{v}_p+t)} \circ \Phi^{\mathbf{i}_-}_{\mathbf{v}_-} \circ \mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}(e^\star).$$

We also let $\sigma^t_{(\mathbf{v}_-,\bar{\mathbf{v}})} = v_1 + \ldots + v_{m-1} + \bar{v}_1 + \ldots + \bar{v}_p + t$. Note that in particular,

$$\Psi^s(\mathbf{v}_-, \bar{\mathbf{u}}, t) = \varphi^{i_m}_{s-\sigma^t_{(\mathbf{v}_-,\bar{\mathbf{u}})}} \circ \Phi^{\mathbf{i}_-}_{\mathbf{v}_-}(x^*) = \psi(\mathbf{v}_-, s - \sigma^t_{(\mathbf{v}_-,\bar{\mathbf{u}})})$$

for all $(\mathbf{v}_-, \bar{\mathbf{u}}, t) \in D^s$. With this property, the next lemma is straightforward :

Lemma 2.3. *For all $k \in \{1, \ldots, m-1\}$, for all $(\mathbf{v}_-, \bar{\mathbf{u}}, t) \in D^s_{m+p}$, one has*

$$\frac{\partial \Psi^s}{\partial v_k}(\mathbf{v}_-, \bar{\mathbf{u}}, t) = -\frac{\partial \psi}{\partial v_m}(\mathbf{v}_-, s - \sigma^t_{(\mathbf{v}_-,\bar{\mathbf{u}})}) + \frac{\partial \psi}{\partial v_k}(\mathbf{v}_-, s - \sigma^t_{(\mathbf{v}_-,\bar{\mathbf{u}})}),$$

and

$$\frac{\partial \Psi^s}{\partial t}(\mathbf{v}_-, \bar{\mathbf{u}}, t) = -\frac{\partial \psi}{\partial v_m}(\mathbf{v}_-, s - \sigma^t_{(\mathbf{v}_-,\bar{\mathbf{u}})}).$$

In particular, setting $s = u_1 + \ldots + u_m + \bar{u}_1 + \ldots + \bar{u}_p$ and $t = 0$, one gets

$$\frac{\partial \Psi^s}{\partial v_k}(\mathbf{u}_-, \bar{\mathbf{u}}, 0) = -\frac{\partial \psi}{\partial v_m}(\mathbf{u}) + \frac{\partial \psi}{\partial v_k}(\mathbf{u}), \tag{2.4}$$

and

$$\frac{\partial \Psi^s}{\partial t}(\mathbf{u}_-, \bar{\mathbf{u}}, 0) = -\frac{\partial \psi}{\partial v_m}(\mathbf{u}). \tag{2.5}$$

Proof of Proposition 2.5

For $s = u_1 + \ldots + u_m + \bar{u}_1 + \ldots + \bar{u}_p$, equalities (2.4) and (2.5) proves that the rank of the family of vectors $(\frac{\partial \Psi^s}{\partial v_1}(\mathbf{u}_-, \bar{\mathbf{u}}, 0), \ldots, \frac{\partial \Psi^s}{\partial v_{m-1}}(\mathbf{u}_-, \bar{\mathbf{u}}, 0), \frac{\partial \Psi^s}{\partial t}(\mathbf{u}_-, \bar{\mathbf{u}}, 0))$ is the same as the family of vectors $(\frac{\partial \psi}{\partial v_k}(\mathbf{u}), 1 \leq k \leq m)$. But since ψ is a submersion at $\mathbf{u}$, this rank is d, showing that Ψ^s is also a submersion at point $(\mathbf{u}_-, \bar{\mathbf{u}}, 0)$. □

We can now pass to the main part of the proof of Proposition 2.1.

Proof of Proposition 2.1

We first construct a function $\bar{\Psi}$ and then verify that it is indeed a submersion. By Proposition 2.5, there exist $s > 0$, $\mathbf{i} = (i_1, \ldots, i_n, i_{n+1}) \in E^{n+1}$ and $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n_+$ such that the map $\Psi : (\mathbf{v}, t) \to \varphi^{i_{n+1}}_{s-\sum v_i - t} \circ \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(e^\star)$ is a submersion at $(\mathbf{u}, 0)$. In the sequel, we denote by $\Psi(\mathbf{v},t)$ the map given by $\Psi(\mathbf{v},t)(x) = \varphi^{i_{n+1}}_{s-\sum v_i - t} \circ \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x)$. We define the map $\overline{\Psi}$ on D^s_{n+N} with values in $\mathbb{R}^d$ by

$$\overline{\Psi}(\mathbf{v},\bar{\mathbf{v}}) \to \varphi^{i_{n+1}}_{s-\sum v_i - \sum \bar{v}_i} \circ \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}} \circ \mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}(e^\star),$$

where $\bar{\mathbf{i}} = (1,2,\ldots,N)$. Then with the previous notation, $\overline{\Psi}(\mathbf{v},\bar{\mathbf{v}}) = \Psi(\mathbf{v},\sum \bar{v}_i) \circ \mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}(e^\star)$. Now, we show that the map $\overline{\Psi}$ is a submersion at $(\mathbf{u},0)$ — here, 0 denotes the zero vector in $\mathbb{R}^N$. For all $1 \le k \le n$,

$$\frac{\partial \overline{\Psi}}{\partial v_k}(\mathbf{v},\bar{\mathbf{v}}) = \frac{\partial \Psi}{\partial v_k}(\mathbf{v},\sum \bar{v}_i) \circ \mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}(e^\star), \tag{2.6}$$

and for all $1 \le k \le N$,

$$\frac{\partial \overline{\Psi}}{\partial \bar{v}_k}(\mathbf{v},\bar{\mathbf{v}}) = \frac{\partial \Psi}{\partial t}(\mathbf{v},\sum \bar{v}_i) \circ \mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}(e^\star) + D\Psi(\mathbf{v},\sum \bar{v}_i)(\mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}(e^\star))\frac{\partial \mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}}{\partial \bar{v}_k}(e^\star). \tag{2.7}$$

Now, since each φ^i_v is the identity at $v=0$ and $\partial_v \varphi^i_v(x) = F^i(\varphi^i_v(x))$, one can easily show that

$$\left.\frac{\partial \mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}}{\partial \bar{v}_k}(e^\star)\right|_{\bar{\mathbf{v}}=0} = F^k(e^\star). \tag{2.8}$$

In particular, since $\mathbf{\Phi}^{\bar{\mathbf{i}}}_{\bar{\mathbf{v}}}(e^\star) = e^\star$ when $\bar{\mathbf{v}} = 0$,

$$\frac{\partial \overline{\Psi}}{\partial \bar{v}_k}(\mathbf{v},0) = \frac{\partial \Psi}{\partial t}(\mathbf{v},0)(e^\star) + D\Psi(\mathbf{v},\sum \bar{v}_i)(e^\star)F^k(e^\star),$$

which, due to condition **(i)** implies that

$$\sum_{k=1}^N \alpha_k \frac{\partial \overline{\Psi}}{\partial \bar{v}_k}(\mathbf{v},0) = \frac{\partial \Psi}{\partial t}(\mathbf{v},0)(e^\star). \tag{2.9}$$

Thus, (2.6) and (2.9) evaluated at $\mathbf{v} = \mathbf{u}$ and $\bar{\mathbf{v}} = 0$ yield

$$\operatorname{rank}\left(\frac{\partial \overline{\Psi}}{\partial v_k}(\mathbf{u},0), \frac{\partial \overline{\Psi}}{\partial \bar{v}_k}(\mathbf{u},0)\right) \ge \operatorname{rank}\left(\frac{\partial \Psi}{\partial v_k}(\mathbf{u},0), \frac{\partial \Psi}{\partial t}(\mathbf{u},0)\right) = d,$$

where the last equality is due to Proposition 2.5. This finishes the proof. □

Chapter 3

Random switching between vector fields having a common zero

In this second chapter, we study the long term behaviour of a PDMP where all the vector fields F^i have a common zero $q \in M$. We show, using stochastic persistence results, 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 q and J is the jump process with rates $(a_{ij}(q))$. 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, $X_t \to q$ exponentially fast when $\Lambda^+ < 0$ and (X, I) converges in distribution at an exponential rate toward a (unique) invariant measure supported by $M \setminus \{q\} \times E$ when $\Lambda^- > 0$. Some applications to certain epidemic models in a fluctuating environment are discussed and illustrate our results.

This joint work with Michel Benaïm has been published in the *Annals of Applied Probability* [BS19]. There are some redundancies with Chapter 1.

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

MSC primary:60J25, 34A37, 37H15, 37A50, 92D30

Contents

3.1 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 $\Psi^i = \{\Psi_t^i\}$ 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

$$\Psi_t^i(M) \subset M$$

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

$$\mathcal{L}g(x,i) = \langle F^i(x), \nabla g^i(x) \rangle + \sum_{j \in E} a_{ij}(x)(g^j(x) - g^i(x)), \tag{3.1}$$

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

$$\frac{dX_t}{dt} = F^{I_t}(X_t), \tag{3.2}$$

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

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

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

where $[0,x] = \{tx, \ t \in [0,1]\}$.

Compactness of M is assumed here for simplicity, but some of the (local) results generalise to noncompact sets. The global results can be extended provided we can control the behaviour of the process near infinity, for instance with a suitable Lyapunov function (see Section 3.3.3).

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 non-linear system is reminiscent of the results obtained by Baxendale [Bax91] and others for Stratanovich stochastic differential equations (see [Bax91] and the references therein, and Hening, Nguyen and Yin [HNY18b] 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 [BH12] and [BLBMZ15] and exposed in Section 1.2.4) and, on the other hand, on some recent results on *stochastic persistence* (Benaïm [Ben18], see Section 1.3), strongly inspired by the seminal works of Schreiber, Hofbauer and their co-authors on *persistence*, first developed for purely deterministic systems (Schreiber [Sch00], Garay and Hofbauer [GH03], Hofbauer and Schreiber [HS04]) and later for certain stochastic systems (Benaïm, Hofbauer and Sandholm [BHS08], Benaïm and Schreiber [BS09], Schreiber, Benaïm and Atchade [SBA11], Schreiber [Sch12], Roth and Schreiber [RS14]).

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 (see [LY76] and Equation (2) in Introduction). 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 [ARBMW14]. 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.

3.1.1 Outline of contents

Section 3.2 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 3.1) 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 3.2) and then for Metzler matrices (Proposition 3.3).

The main results of the paper are stated in Section 3.3.

- 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 3.4 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 [LY76] (or more generally sub homogeneous cooperative systems in the sense of Hirsch [Hir94]). 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 3.10).

Sections 3.5 and 3.6 are devoted to the proofs of Theorems 3.1, 3.2, 3.3, 3.4 and 3.10. The proofs of certain results stated in Section 3.2 are given in appendix (Section 3.7) for convenience.

3.1.2 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 In addition to the notation introduced in Chapter 1, we will use the following ones. 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 càdlàg paths. Given such a process and $\mu \in \mathcal{P}(\mathcal{X})$ we let $\mathbb{P}^{\tilde{Z}}_\mu$ denote the law of $\tilde{Z}$ on the Skorokhod space $\mathbb{D}(\mathbb{R}^+, \mathcal{X})$ when $\tilde{Z}_0$ has law μ. As usual, $\mathbb{P}^{\tilde{Z}}_z$ stands for $\mathbb{P}^{\tilde{Z}}_{\delta_z}$ for all $z \in \mathcal{X}$. We let $(P^{\tilde{Z}}_t)_{t \geq 0}$ be the semigroup of $\tilde{Z}$. We let $\mathcal{P}^{\tilde{Z}}_{inv} \subset \mathcal{P}(\mathcal{X})$ denote the (possibly empty) set of invariant probabilities of $\tilde{Z}$ and $\mathcal{P}^{\tilde{Z}}_{erg} \subset \mathcal{P}^{\tilde{Z}}_{inv}$ the subset of ergodic probabilities. Recall that $\mathcal{P}^{\tilde{Z}}_{erg}$ can also be defined as the set of extremal points of $\mathcal{P}^{\tilde{Z}}_{inv}$.

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

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

We refer the reader to Meyn and Tweedie ([MT09], chapter 17) for a proof and more details.

3.2 The Linearised 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

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

where δ is like in condition $C2$. Here, $\overline{B}$ stands for the closure of B.

Remark 3.1. *One can check that the definition of C_M does not depend on the choice of δ, provided δ satisfies condition $C2$.*

Lemma 3.1. *For all $t \geq 0$, $e^{tA^i} C_M \subset C_M$.*

Proof We set $D_M = \{tx : t \geq 0, x \in M, \|x\| \leq \delta\}$ and first prove that $e^{tA^i} D_M \subset C_M$. The lemma will be then induced by continuity of e^{tA^i}. Let $x \in D_M$. For ε small enough, by definition of D_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 $D\Psi_t^i(0)x = e^{tA^i} x \in C_M$. □

Define the *linearised 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

$$\begin{cases} \frac{dY_t}{dt} = A^{J_t} Y_t \\ Y_0 = y, \end{cases} \tag{3.3}$$

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 (3.3) can be rewritten as

$$\begin{cases} \frac{d\Theta_t}{dt} = G^{J_t}(\Theta_t) \\ \frac{d\rho_t}{dt} = \langle A^{J_t} \Theta_t, \Theta_t \rangle \rho_t, \end{cases} \tag{3.4}$$

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

$$G^i(\theta) = A^i \theta - \langle A^i \theta, \theta \rangle \theta. \tag{3.5}$$

Remark 3.2. For stochastic differential equations, the idea of introducing, this polar decomposition goes back to Hasminskii [Has60] and has proved to be a fundamental tool for analyzing linear stochastic differential equations (see e.g [Bax91]), linear random dynamical systems (see e.g chapter 6 of Arnold [Arn98]) and more recently certain linear PDMPs in [BLBMZ14], [LMR14] or [Lag16].

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 Propositions 1.9 and 1.6), $\mathcal{P}_{inv}^{(\Theta,J)}$ is a nonempty compact (for the topology of weak* convergence) subset of $\mathcal{P}(S^{d-1} \cap C_M \times E)$.

3.2.1 Average growth rates

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

$$\Lambda(\mu) = \int \langle A^i\theta, \theta\rangle \mu(d\theta di) = \sum_{i \in E} \int_{S^{d-1}\cap C_M} \langle A^i\theta, \theta\rangle \mu^i(d\theta), \tag{3.6}$$

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 (3.4) 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

$$\Lambda^- = \inf\{\Lambda(\mu) \ : \mu \in \mathcal{P}_{erg}^{(\Theta,J)}\} \text{ and } \Lambda^+ = \sup\{\Lambda(\mu) \ : \mu \in \mathcal{P}_{erg}^{(\Theta,J)}\}. \tag{3.7}$$

The following rough estimate is a direct consequence of (3.6). Recall that $p = (p_i)_{i\in E}$ is the invariant probability of J.

Lemma 3.2.

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

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 (see Theorem 1.17 and Proposition 1.15).

3.2.2 Relation with Lyapunov exponents

Set $\Omega = \mathbb{D}(\mathbb{R}_+, E)$ 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 $\varphi(0, \omega)y = y$.

Recall that in Section 1.4, we show that φ is an ergodic 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 Proposition 1.15). 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 $\mathbb{P}_p^J(\tilde{\Omega}) = 1$, 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

$$\lim_{t \to \infty} \frac{1}{t} \log \|\varphi(t, \omega)y\| = \lambda_i \tag{3.8}$$

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

Proposition 3.1. *For all $\mu \in \mathcal{P}_{erg}^{(\Theta,J)}$*

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

If furthermore C_M has non empty interior, then

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

Remark 3.3. The second part of the proposition has already been proven by Crauel [Cra84, Theorem 2.1 and Corollary 2.2] in a more general setting. We adapt the arguments of his proof for our specific case.

Proof Let $\mu \in \mathcal{P}_{erg}^{(\Theta,J)}$. Then, $\mathbb{P}_\mu^{(\Theta,J)}$ 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 (3.3), (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 $(\theta, \omega) \in \mathcal{B}$

$$\lim_{t \to \infty} \frac{1}{t} \log(\|\varphi(t, \omega)\theta\|) = \Lambda(\mu) \tag{3.9}$$

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

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 $\mathbb{P}_\mu^{(\Theta_0,J)}(S^{d-1}\cap C_M \times \tilde{\Omega}) = \mathbb{P}_\mu^J(\tilde{\Omega}) = 1$. Hence $\mathbb{P}_\mu^{(\Theta_0,J)}(\tilde{\mathcal{B}}) = 1$ and for all $(\theta,\omega) \in \tilde{\mathcal{B}}$ the left hand side of equality (3.9) equals λ_i for some i.

It remains to show that $\lambda_1 = \Lambda^+$. For every ω in the set $\tilde{\Omega}$ given by the multiplicative ergodic theorem, and for all $\theta \in S^{d-1}\cap C_M$, define

$$\lambda(\theta,\omega) = \lim_{t\to\infty}\frac{1}{t}\log(\|\varphi(t,\omega)\theta\|) = \lim_{t\to\infty}\frac{1}{t}\int_0^t \langle A^{\omega_s}\Theta_s^\theta(\omega), \Theta_s^\theta(\omega)\rangle ds,$$

where

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

By (4.9), we have $\lambda(\theta,\omega) = \lambda_1$ for all $\theta \in V_1(\omega)\setminus V_2(\omega)\cap S^{d-1}\cap C_M$. Let ν denote the normalised Lebesgue measure on $S^{d-1}\cap C_M$. Because $V_2(\omega)$ is at most an hyperplane and C_M has non empty interior, we get that $\int\lambda(\theta,\omega)d\nu(\theta) = \lambda_1$ for all $\omega\in\tilde{\Omega}$. In particular,

$$\int_\Omega\int_{S^{d-1}\cap C_M}\lambda(\theta,\omega)d\nu(\theta)d\mathbb{P}_p^J(\omega) = \lambda_1. \tag{3.10}$$

Moreover, because $|\langle A^i\theta,\theta\rangle| \leq \max\|A^i\|$, dominated convergence and (3.10) imply that

$$\lambda_1 = \lim_{t\to\infty}\frac{1}{t}\int_\Omega\int_{S^{d-1}\cap C_M}\int_0^t\langle A^{\omega_s}\Theta_s^\theta(\omega),\Theta_s^\theta(\omega)\rangle ds d\nu(\theta)d\mathbb{P}_p^J(\omega) \tag{3.11}$$

Now for all $t > 0$, define the probability on $S^{d-1}\cap C_M\times E$

$$\mu_t = \frac{1}{t}\int_0^t(\nu\otimes p)P_s^{(\Theta,J)}ds. \tag{3.12}$$

By compactness of $S^{d-1}\cap C_M\times E$, $(\mu_t)_{t\geq 0}$ is tight, and by Feller property of (Θ, J), every weak limit points of μ_t belongs to $\mathcal{P}_{inv}^{(\Theta,J)}(S^{d-1}\cap C_M\times E)$. Let μ be such a limit point, and (t_n) such that $\mu_{t_n}\to\mu$. Setting $f(\theta,i) = \langle A^i\theta,\theta\rangle$, one has $\mu_{t_n}f\to\mu f = \Lambda(\mu)$. Now (3.9), (3.11) and Fubini Theorem imply that $\lambda_1 = \lim\mu_{t_n}f = \Lambda(\mu)$, which concludes the proof. □

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 Theorem 1.17 for details). A consequence of Proposition 3.1 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 3.2 and 3.3 and Theorems 3.2 and 3.3).

Corollary 3.1.

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

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 by Corollary 1.3, the left-hand side converges $\mathbb{P}_p^J$ almost surely, as $t \to \infty$, toward $\sum_{i=1}^{\tilde{d}} d_i \lambda_i$. $\square$

Remark 3.4. If the matrices A^i are *Metzler*, meaning that they have off diagonal non-negative entries, a result due to Mierczyński ([Mie15], Theorem 1.3) allows to improve the lower bound given in Corollary 3.1 We will use this estimate in section 3.4, example 3.7.

Remark 3.5. Note that in general

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

Here is a simple example based on [BLBMZ14]. 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 [BLBMZ14], Example 1.3 and Example 1.1 in Chapter 1). 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 3.1, 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 [BLBMZ14] (see Theorem 1.6 in Chapter 1), that $\Lambda^+ = \Lambda^- > 0$ for β sufficiently large. Hence (for large β)

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

3.2.3 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 3.5). 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 3.1 (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. Recall the Definitions 1.15 and 1.16 given in the Chapter 1.

Proposition 3.2. *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 $\{G^i\}$-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 Proposition 1.13, 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. $\square$

Example 3.2 (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 3.2 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 3.2. 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 3.2 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 3.4 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 [AGD94]. However, to handle the weaker assumption (iii), the proof needs to be adapted, but relies on the same ideas. Details are given in Section 3.7. Recall (see remark 3.4) that a *Metzler* matrix is a matrix with nonnegative off-diagonal entries. We say that such a matrix is *irreducible* if adding a sufficiently large multiple of the identity, the obtained matrix is a non-negative irreducible matrix in the usual sense.

Proposition 3.3. *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*

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

3.2.4 Average growth rate under frequent switching

The definition of average growth rates (see equations (3.6) and (3.7)) 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 3.3. *Assume the switching rates are constant and depend on a small parameter* $\varepsilon : a^{\varepsilon}_{i,j} = a_{i,j}/\varepsilon$ *where* $(a_{i,j})$ *is an irreducible matrix with invariant probability* p*. Denote by* $(\Theta^\varepsilon, J^\varepsilon)$ *the associated PDMP given by* (3.4)*, 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 3.7. An immediate corollary is :

Corollary 3.2. *With the hypotheses of Lemma 3.3, 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*

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

3.3 The non linear system : Main results

3.3.1 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 $\{F^i\}$-accessible from M, then for all $x \in M$ and $i \in E$

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

3.3.2 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 $i \in E$,

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

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 $p \in M^*$ $\{F^i\}$-accessible from M^* at which the weak bracket condition holds. Then*

(i) *The set $\mathcal{P}^Z_{inv} \cap \mathcal{P}(M^* \times E)$ reduces to a single element, denoted Π;*

(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 of the process $(Z_t)_{t\geq 0}$, the weak bracket condition needs to be strengthened. Recall that 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 conditions of the preceding theorem, assume furthermore that one the two following holds :*

(i) *The weak bracket condition is strengthened to the strong bracket condition; or*

(ii) *There exist $\alpha_1, \dots, \alpha_N \in \mathbb{R}$ with $\sum \alpha_i = 1$ and a point $e^\star \in M^*$ $\{F^i\}$-accessible from M^* such that $\sum \alpha_i F^i(e^\star) = 0$.*

Then there exist $\kappa, \theta > 0$ such that for all $x \in M^*$ and $i \in E$,

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

3.3.3 The noncompact case

We briefly discus here the situation where M is not compact. First, note that all the results given in section 3.2 still hold, because they only deal with the linearised system. Next, local statements remain true without additional assumption by a localisation argument. Namely :

Theorem 3.5.

(i) *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 $i \in E$*

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

(ii) *Assume $\Lambda^- > 0$. Then there exist $\varepsilon > 0$, $b > 1$ and $c > 0$ such that for all $x \in M^*$ and $i \in E$,*

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

Point **(ii)** is rigorously proven in Section 4.3 of Chapter 4. To extend the global results stated above , we make the additional assumption that the jumps rates are bounded and that there exists a Lyapunov function, controlling the behaviour of the process at infinity.

Assumption 3.1. *The jumps rate are bounded :*

$$\sup_{x\in M}\max_{i,j} a_{ij}(x) < \infty.$$

For a function $f : M \times E \to \mathbb{R}$, we denote by Γf the function defined by :

$$\Gamma f(x,i) = \sum_{j\in E} a_{ij}(x)\left(f(x,j) - f(x,i)\right)^2.$$

We also let C^1_c denote the space of functions $f : M \times E \to \mathbb{R}$ that are constant outside a compact set and C^1 in the first variable.

Assumption 3.2. *There exists a continuous function $W : M\times E \to \mathbb{R}_+$ with $\lim_{\|x\|\to\infty} W(x,i) = \infty$, a continuous function $LW : M \times E \to \mathbb{R}_+$, $\alpha > 0$ and $C \geq 0$ such that*

- ***(i)*** *For every compact set $K \subset M$, there exists $W_K \in C^1_c$ such that*
  - ***(a)*** *$W|_K = W_K|_K$ and $\mathcal{L}W_K|_K = LW|_K$,*
  - ***(b)*** *For all $x \in M$, $\sup\{P_t(\Gamma W_K),\quad t \geq 0,\quad K\quad compact\} < \infty$*
- ***(ii)***

$$LW \leq -\alpha W + C.$$

Theorem 3.6. *Under Hypotheses 3.2 and 3.1, Theorems 3.1, 3.2 and 3.3 are still valid. Moreover, Theorem 3.4 is true, but with the following estimate :*

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

Example 3.3. We consider a random switching between two linear systems given by 2×2 Metzler matrices A^0 and A^1, with transition rate $a_{i,1-i}(x)$. We assume that A^0 has two distinct positive eigenvalues $\lambda_1 > \lambda_2$ and is irreducible, whereas A^1 is of the form

$$A^1 = \begin{pmatrix} -c & 0 \\ 0 & -d \end{pmatrix},$$

with $0 < c < d$. Since the eigenvalues of A^0 are positive, there is no invariant compact set for Ψ^0, nor for the PDMP. Moreover, A^0 and A^1 being Metzler, $M = \mathbb{R}^2_+$ is positively invariant for $(X_t)_{t\geq 0}$. If the jump rates were constant in x, the process would either converge to 0 or to infinity. To ensure positive recurrence on M^*, we assume that the transition rates are such that, near the origin, I_t spends more time in state 0 :

$$a_{10}(0) - \frac{d}{\lambda_2}a_{01}(0) > 0; \tag{3.13}$$

While near infinity, it spends more time in state 1 :

$$\limsup_{\|x\|\to\infty}\left(a_{10}(x) - \frac{c}{\lambda_1}a_{01}(x)\right) < 0. \tag{3.14}$$

More precisely, we have the following :

Proposition 3.4. *Assume that the jumps rates are bounded and that conditions* (3.13) *and* (3.14) *hold. Then there exists a unique invariant probability* $\Pi \in \mathcal{P}(M^* \times E)$ *and there exists* $\kappa, \theta, q > 0$ *such that for all* $x \in M^*$ *and* $i \in E$,

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

Proof By Theorem 3.3, $\Lambda^+ = \Lambda^- := \Lambda$, and by Corollary 3.1,

$$\Lambda \geq \frac{1}{2}(p_0 \operatorname{Tr}(A^0) + p_1 \operatorname{Tr}(A^1) \geq \lambda_2 p_0 - dp_1.$$

Moreover, it is easy to check that $p_0 = \frac{a_{10}(0)}{a_{10}(0)+a_{01}(0)}$ and $p_1 = \frac{a_{01}(0)}{a_{10}(0)+a_{01}(0)}$. Hence, if $a_{10}(0) > \frac{d}{\lambda_2} a_{01}(0)$, then $\Lambda > 0$. Now we show that we can construct a Lyapunov function at infinity. Let $q > 0$ and $\beta_0, \beta_1 > 0$ and define, for all $(x, i) \in M \times E$, $W_q(x, i) = \beta_i \|x\|^q$. Formally, we have

$$\mathcal{L}W_q(x, i) = q\beta_i \langle A_i x, x\rangle \|x\|^{p-2} + a_{i,1-i}(x)(\beta_{1-i} - \beta_i)\|x\|^q.$$

By assumption on A^0 and A^1, $\langle A_0 x, x\rangle \leq \lambda_1 \|x\|^2$ and $\langle A_1 x, x\rangle \leq -c\|x\|^2$. Hence,

$$\mathcal{L}W_q(x, i) \leq \left(-\alpha(i)q\beta_i + a_{i,1-i}(x)(\beta_{1-i} - \beta_i)\right) \|x\|^q,$$

where $\alpha(0) = -\lambda_1$ and $\alpha(1) = c$. First we prove that we can choose β_0 and β_1 such that W_q satisfies point **(ii)** of Hypothesis 3.2 for all q small enough. Then we prove that we can choose q such that point **(i-b)** holds. By assumption (3.14), there exists $\varepsilon > 0$ and $K > 0$ such that, for all $x \in M$ with $\|x\| \geq K$, $a_{10}(x) \leq \frac{c}{\lambda_1} a_{01}(x) - \varepsilon$. This implies that, for q small enough, there exists α_q such that $a_{10}(x)(\frac{\alpha_q}{\lambda_1} + q) - (\frac{c}{\lambda_1} - \frac{\alpha_q}{\lambda_1})a_{01}(x) - q\alpha_q + cq^2 \leq 0$, which yields

$$\sup_{\|x\|\geq K} \frac{a_{01}(x) + \alpha_q}{a_{01}(x) - \lambda_1 q} \leq \inf_{\|x\|\geq K} \frac{-\alpha_q + cq}{a_{10}(x)} + 1.$$

Now we choose $\beta_1 = 1$ and β_0 such that

$$\sup_{\|x\|\geq K} \frac{a_{01}(x) + \alpha_p}{a_{01}(x) - \lambda_1 q} \leq \beta_0 \leq \inf_{\|x\|\geq K} \frac{-\alpha_q + cq}{a_{10}(x)} + 1.$$

Thus, for $\|x\| \geq K$, $-\alpha(i)q\beta_i + a_{i,1-i}(x)(\beta_{1-i} - \beta_i) \leq -\alpha_q$. In particular, for all for $\|x\| \geq K$, $\mathcal{L}W_q(x, i) \leq -\alpha_q W_q(x, i)$. Since $\mathcal{L}W_q$ is bounded for $\|x\| \leq K$, then $\mathcal{L}W_q \leq -\alpha_q W_q + C$ for some constant $C > 0$ (depending on $q > 0$). This has the consequence (see [Ben18, Theorem 2.1]) that for all $t \geq 0$,

$$P_t W_q \leq e^{-\alpha_q t}\left(W_q - \frac{C}{\alpha_q}\right) + \frac{C}{\alpha_q}. \tag{3.15}$$

The computation of Γ gives

$$\Gamma W_q(x, i) = a_{i,1-i}(x)(\beta_0 - \beta_1)^2 \|x\|^{2q},$$

hence

$$\Gamma W_q \leq \tilde{C}_q W_{2q}$$

for some constant $\tilde{C}_q > 0$. Hence, choosing p small enough so that (3.15) holds for $2q$, one has

$$\sup_{t\geq 0} P_t\left(\Gamma W_q\right) \leq \tilde{C}_q \sup_{t\geq 0} P_t W_{2q} \leq W_{2q},$$

which proves **(i-b)**. It remains to show that there exist accessible points at which the strong bracket condition holds. Set $F^0(x) = A^0x$ and $F^1(x) = A^1x$ the vector fields associated to A^0 and A^1. There exist $\alpha, \beta, \gamma, \delta$, with $\beta, \gamma > 0$ such that $F^0(x,y) = (\alpha x + \beta y, \gamma x + \delta y)$. Straightforward computations show that

$$\det(F^0 - F^1, [F^0, F^1])(x,y) = (d-c)(2\beta\gamma xy + \beta(d+\delta)y^2 + \gamma(\alpha + c)x^2).$$

Since $\beta, \gamma > 0$, this polynomial is non identically null. To conclude, we prove that there exists an open set of accessible points. Let $v \in \mathbb{R}^2_{++}$ be the Perron eigenvector associated with A^0. We claim that $\mathbb{R}_+ v$ and therefore $\gamma_1^+(\mathbb{R}_+ v) = \overline{\cup_{t\geq 0}\Psi_t^1(\mathbb{R}_+ v)}$ are accessible. One can check that for all $y \in \mathbb{R}_+ v$ and all $\varepsilon > 0$, there exists $\eta > 0$ such that for all $x \in M^*$ with $\|x\| < \eta$, there exists $t \geq 0$ such that $\|\Psi_t^0(x) - y\| < \varepsilon$. Since 0 is accessible following F^1, this makes y accessible. Hence, $\gamma_1^+(\mathbb{R}_+ v)$ is accessible and Theorem 3.6 applies. $\square$

3.4 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 [LY76], 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

$$\frac{dx_i}{dt} = (1 - x_i)(\sum_{j=1}^{d} C_{ij}x_j) - D_i x_i\,,\ i = 1, \dots d, \tag{3.16}$$

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, (3.16) can be derived from a finite population model, in the limit of an infinite population (see Benaïm and Hirsch [BH99]).

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 [LY76] prove the following result:

Theorem 3.7 (Lajmanovich and Yorke, [LY76])**.** *Let* $A = C - \mathsf{diag}(D)$.

If $\lambda(A) \leq 0$, 0 *is globally asymptotically stable for the semiflow induced by (3.16) 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 [BH99] (especially Section 3), we call a map $F : [0,1]^d \mapsto \mathbb{R}^d$ an *epidemic vector field* if it is continuously differentiable[1] and 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 (3.16)) 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 ([BH99], 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 [Hir94]).

Theorem 3.8. *Let F be an epidemic vector field and $\Psi = \{\Psi_t\}_{t\geq 0}$ the induced semiflow on $[0,1]^d$. Then*

(i) *(Hirsch, [Hir94]) Either 0 is globally asymptotically stable for Ψ; or there exists another equilibrium $x^* \in]0,1[^d$ whose basin of attraction is $[0,1]^d \setminus \{0\}$.*

(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 [Hir94], 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. $\square$

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

3.4.1 Fluctuating environment

We consider a PDMP $Z = (X, I)$ as defined in Section 3.1, 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 3.3, we let $\lambda_1 = \Lambda^+ = \Lambda^-$ denote the top Lyapunov exponent of the linearized system.

Theorem 3.9. *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 3.2.2) that Ω stands for $\mathbb{D}(\mathbb{R}^+, E)$. 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 [Hir94], 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 $\|x\| > \varepsilon$

$$\Psi(t, \omega)(x) \leq \frac{\|x\|}{\varepsilon} \Psi(t, \omega)(\frac{\varepsilon}{\|x\|} x). \tag{3.17}$$

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 $x \in B(0, \varepsilon)$

$$\mathbb{P}^Z_{x,i}(\limsup_{t \to \infty} \frac{\log(\|X_t\|)}{t} \leq \lambda_1) = \mathbb{P}^J_i(\limsup_{t \to \infty} \frac{\log(\|\Psi(t, \omega)(x)\|)}{t} \leq \lambda_1) \geq \eta. \tag{3.18}$$

Combined with (3.17), this proves that (3.18) 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 3.8, that 0 is $\{F^i\}$-accessible from M, and the result follows from Theorem 3.1. □

Remark 3.6. 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 3.7. Case (a) (and its proof) can be related with the results obtained by Chueshov in [Chu02], for SIS models with random coefficients (see [Chu02, 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 3.4 (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

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

One can easily check that

$$\lambda(A^0) = \lambda(A^1) = \sqrt{5} - 2 > 0,$$

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

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

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 3.2, this implies that $\lambda_1 < 0$ provided β is sufficiently large. Thus the conclusion of Theorem 3.9 holds.

Example 3.5 (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.

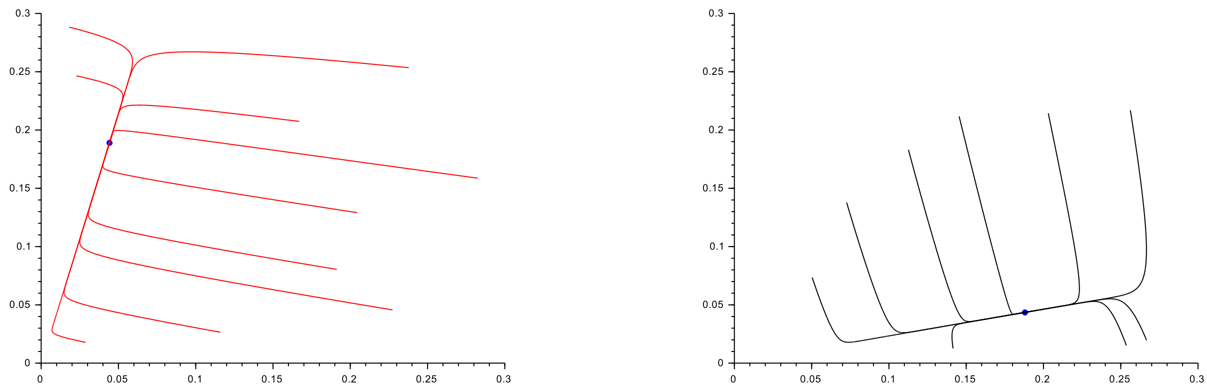


Figure 3.1: Example 3.4, phase portrait of F^0 and F^1

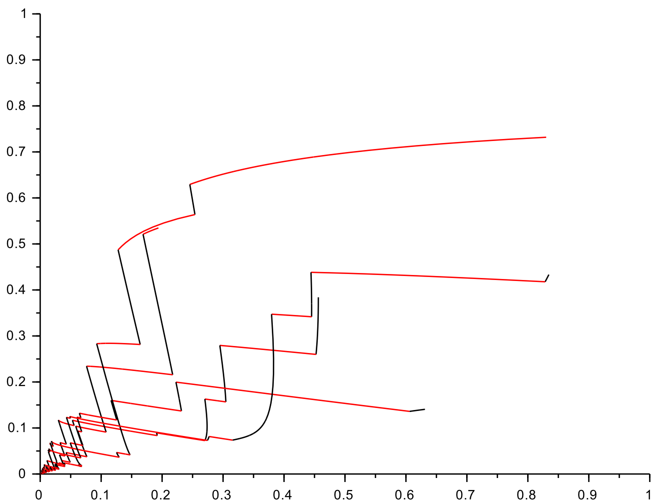


Figure 3.2: Example 3.4, some trajectories of X_t for $\beta = 20$

With the notation of Example 3.4, 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,

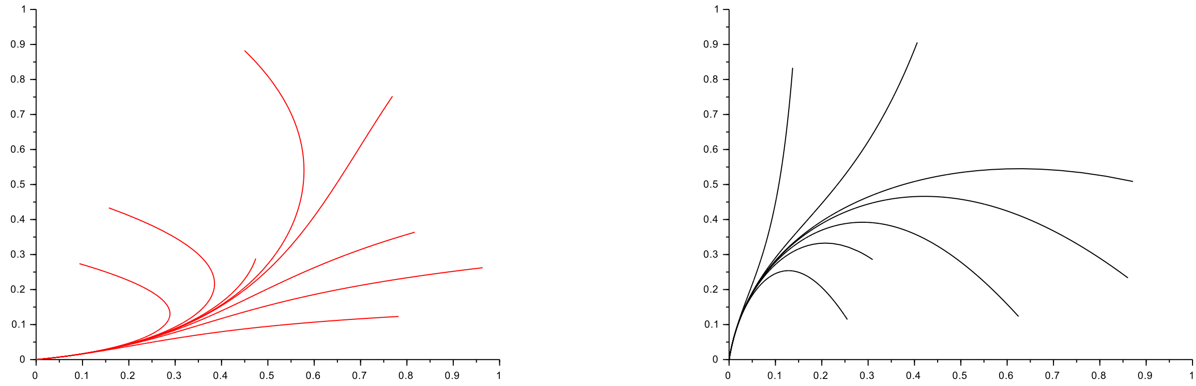


Figure 3.3: Example 3.5, Phase portrait of F^0 and F^1

for β sufficiently large, this implies by Corollary 3.2 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 [BLBMZ15, Proposition 3.1] (see Proposition 1.12 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.

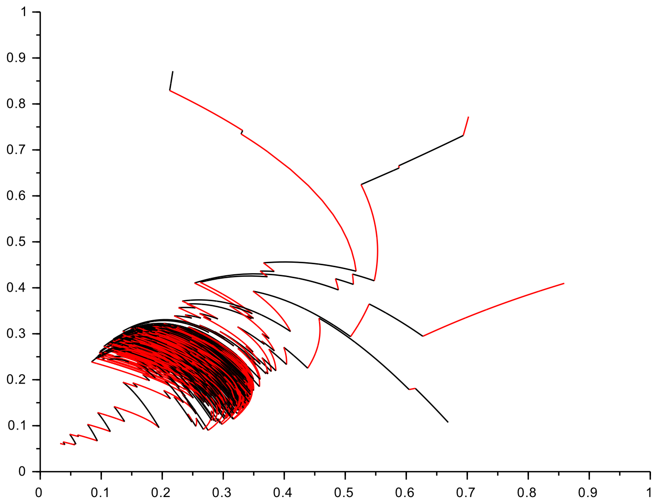


Figure 3.4: Example 3.5, some trajectories of (X_t) for $\beta = 20$

It turns out that the previous example can be generalised in the following way.

Lemma 3.4. *Assume that F^0 and F^1 are two epidemic vector fields in dimension 2 such that*

1. $\lambda(A^0) < 0$ *and* $\lambda(A^1) < 0$,
2. *There exists* $s \in (0,1)$ *such that* $\lambda(A^s) > 0$, *where* $A^s = sA^1 + (1-s)A^0$.

*Then there exists an $\{F^i\}$ - accessible point at which the weak bracket condition holds and condition **(ii)** of Theorem 3.4 is satisfied.*

In particular, Theorem 3.4 implies that there is convergence in total variation to a unique invariant probability measure provided $\lambda_1 > 0$. This happens for example with switching rates of the form

$$a_{0,1} = s\beta, a_{1,0} = (1-s)\beta, a_{0,0} = a_{1,1} = 0.$$

for β large enough (by Corollary 3.2.)

Proof of Lemma 3.4 It is readily seen that for $s \in (0,1)$, the vector field $F^s = sF^1 + (1-s)F^0$ is also an epidemic vector field. As a consequence, since there exists $s \in (0,1)$ such that $\lambda(A^s) > 0$, Theorem 3.8 implies that there exists some point $x_s^* \in (0,1)^2$ such that $F^s(x_s^*) = 0$. In particular, condition **(ii)** of Theorem 3.4 is satisfied. Moreover, since $\lambda(A^0) < 0$ and $\lambda(A^1) < 0$, the first part of Theorem 3.8 implies that neither F^0 nor F^1 can vanish at x_s^*. In particular, $F^0(x_s^*)$ and $F^1(x_s^*)$ are collinear and of opposite direction. For $k \in \{0,1\}$ let $\gamma^k(x_s^*)$ denote the positive orbit of x_s^* under F^k. Due to the first part of Theorem 3.8, $\gamma^0(x_s^*)$ is a curve linking x_s^* and 0. To obtain a contradiction, assume that the weak bracket condition holds nowhere on $\gamma^0(x_s^*)$. Then F^0 and F^1 are collinear and of opposite direction on $\gamma^0(x_s^*)$. We have for all $x \in \gamma^0(x_s^*)$ that $x_s^* \in \gamma^1(x)$, meaning that for all $\varepsilon > 0$, one can find x with $\|x\| < \varepsilon$ and $t > 0$ such that $\|\varphi_t^1(x)\| = \|x_s^*\|$. This is in contradiction with the fact that 0 is a globally asymptotically stable equilibrium for F^1, hence the weak bracket condition holds at some point x' on $\gamma^0(x_s^*)$. Since x_s^* is $\{F^i\}$-accessible, the point x' is also $\{F^i\}$- accessible. $\square$

Remark 3.8. 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 [LMR14] (see Example 1.2) and more recently in [Lag16].

Now, if we restrain ourselves to Metzler matrices, a result from Gurvits, Shorten and Mason ([GSM07, 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.

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 [FMC09]. 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 [FMC09] 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 3.5) 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 3.6 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.

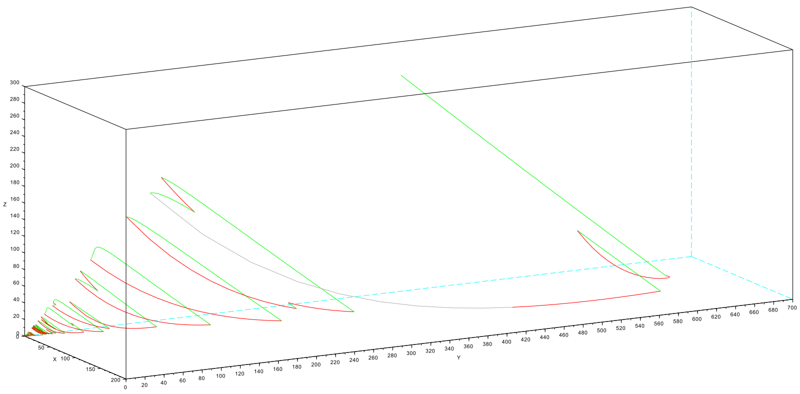


Figure 3.5: Simulation of Y_t for $\beta = 10$.

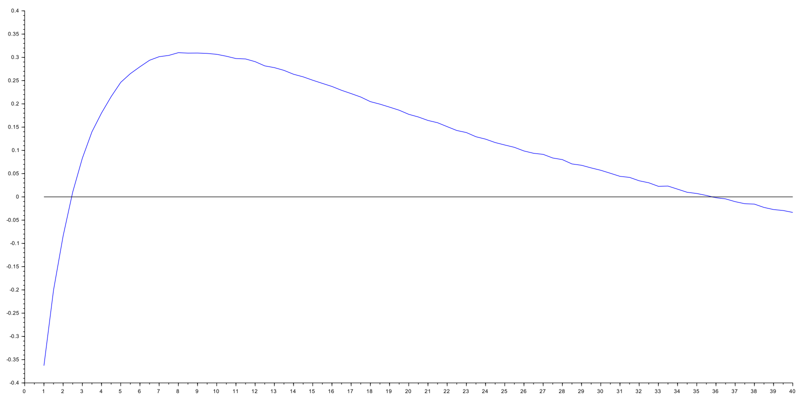


Figure 3.6: Approximation of $\lambda_1(\beta)$ by Monte-Carlo method

Example 3.6 (Fluctuations may promote infection, continued)**.** Remark 3.8 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 $F^t = (1-t)F^0 + tF^1$;

(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 3.6) β is such that $\lambda_1 > 0$, then by Theorem 3.10 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 3.7 and 3.8 illustrate this persistence of the infection. In figure 3.8 , we have plotted $\|X_t\|_1 = X_t^1 + X_t^2 + X_t^3$.

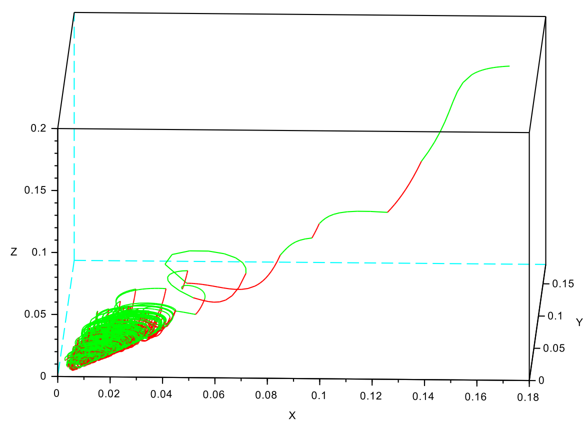


Figure 3.7: Example 3.6 : Simulation of X_t for $\beta = 10$.

3.4.2 Exponential convergence without bracket condition

Throughout this section, we assume that the vector fields F^i are epidemic and that the jump rates are constant. Recall (see proof of Theorem 3.9) 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 expect uniqueness of the invariant measure on $[0,1]^d \setminus \{0\} \times E$ and convergence in law of (Z_t) towards it.

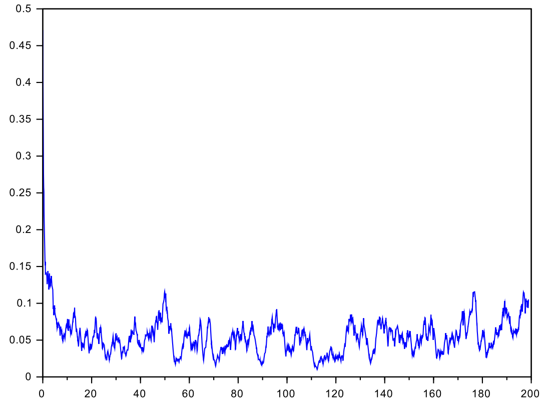


Figure 3.8: Example 3.6 : Simulation of $\|X_t\|_1$ for $\beta = 10$.

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 [HMS11]. 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 3.7. 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 3.3 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 3.4 due to Mierczyński ([Mie15, 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 3.2 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 $i \in E$

$$\mathbb{P}^Z_{x,i}(\limsup_{t \to \infty} \frac{1}{t} \log(\|X_t - x^*\|) \leq \frac{\lambda_1(x^*)}{2}) \geq \eta. \tag{3.19}$$

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

$$\mathbb{P}^Z_{x,i}(\tau < \infty) = 1, \tag{3.20}$$

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 $\|x\| \geq \varepsilon$,

$$\mathbb{P}^Z_{x,i}(Z_t \in \mathcal{U} \times E) \geq c. \tag{3.21}$$

Combining (3.19), (3.20), (3.21) 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 [BL16, 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 3.10. *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 $\mu, \nu \in \mathcal{P}(\mathcal{Y})$,*

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

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

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

3.5 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 [Ben18] and exposed in Section 1.3 of Chapter 1.

In order to apply the results of Section 1.3, 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

$$\begin{cases} \frac{d\rho_t}{dt} = \langle \Theta_t, \tilde{F}^{I_t}(\rho_t, \Theta_t)\rangle \rho_t \\ \frac{d\Theta_t}{dt} = \tilde{F}^{I_t}(\rho_t, \Theta_t) - \langle \Theta_t, \tilde{F}^{I_t}(\rho_t, \Theta_t)\rangle \Theta_t \\ \mathsf{P}(I_{t+s} = j | \mathcal{F}_t) = a_{ij}(\rho_t \Theta_t)s + o(s) \text{ for } i \neq j \text{ on } \{I_t = i\} \end{cases} \tag{3.22}$$

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, (3.22) 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

$$\tilde{\mathcal{L}}f(\rho, \theta, i) = \frac{\partial f^i}{\partial \rho}(\rho, \theta)\langle \theta, \tilde{F}^i(\rho, \theta)\rangle \rho + \langle \nabla_\theta f^i(\rho, \theta), \tilde{G}^i(\rho, \theta)\rangle + \sum_{j \in E} a_{ij}(\rho\theta)(f^j(\rho, \theta) - f^i(\rho, \theta)), \tag{3.23}$$

where $\tilde{G}^i(\rho, \theta) = \tilde{F}^i(\rho, \theta) - \langle \theta, \tilde{F}^i(\rho, \theta)\rangle\theta$. By Proposition 1.9 , $\tilde{Z}$ is Feller. Moreover by equation (3.22), Assumption 1.3 is verified. The following lemma gives V and H that fulfil Assumption 1.5.

Lemma 3.5. *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 Assumption 1.5.*

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)**, **(d)** and **(e)** are clearly satisfied (recall that since M is compact, one can choose $\tilde{W} = 0$ in Assumption 1.4). □

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

Lemma 3.6. *With the notation of the previous sections,*

$$\Lambda^+(H) = \Lambda^+ \quad \textit{and} \quad \Lambda^-(H) = \Lambda^-.$$

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 5.2.2. Now, $\langle A^i\theta, \theta\rangle = -H(0, \theta, i)$, and the result easily follows from the definitions of $\Lambda^{+/-}, \Lambda^{+/-}(H)$
□

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 3.6 $\tilde{Z}$ is H - nonpersistent. Theorem 1.14 **(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 Proposition 1.10, 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 1.14 **(ii)** proves the second assertion of Theorem 3.1. $\square$

To show the other theorems, we use the following lemma for which the proof is omitted. Here, φ denotes Ψ^{-1}.

Lemma 3.7. *The map*

$$\begin{array}{ccc} \mathcal{P}^{\tilde{Z}}_{inv}(\mathcal{X}_+) & \longrightarrow & \mathcal{P}^{Z}_{inv}(M^* \times E) \\ \Pi & \longmapsto & \Pi \circ \varphi^{-1} \end{array}$$

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

$$\Pi^{x,i}_t = \tilde{\Pi}^{\Psi(x,i)}_t \circ \varphi^{-1}.$$

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

Proof of Theorem 3.2 Here we assume $\Lambda^- > 0$, thus by Lemma 3.6 $\tilde{Z}$ is H - persistent. Then Theorem 1.13 **(i)** and Lemma 3.7 imply **(i)** of Theorem 3.2. Moreover, by Theorem 1.13 **(ii)**, we have for some positive θ, K, T

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

Let $\tilde{\mu} \in \mathcal{P}^{\tilde{Z}}_{inv}(\mathcal{X}_+)$ 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

$$\tilde{\mu}\tilde{W} \leq \frac{K}{1-\rho}. \tag{3.24}$$

Now let $\mu \in \mathcal{P}^{Z}_{inv}(M^* \times E)$ 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 3.7, $\mu \circ \Psi^{-1} \in \mathcal{P}^{\tilde{Z}}_{inv}(\mathcal{X}_+)$, and because $W \circ \Psi^{-1} = \tilde{W}$, (3.24) proves **(ii)** of Theorem 3.2. Point **(iii)** is immediate from **(iii)** of Theorem 1.13. $\square$

Proof of Theorem 3.3 By Theorem 3.2, $\mathcal{P}^{Z}_{inv}(M^* \times E)$ is non-empty. So the weak bracket condition implies by Proposition 1.13 uniqueness of Π and the absolute continuity. Moreover, for all $(x,i) \in M^* \times E$, $(\Pi^{x,i}_t)_{t \geq 0}$ is tight and admits a unique limit point Π, so that $\Pi^{x,i}_t$ converges almost surely to Π. $\square$

Proof of Theorem 3.4 Assume that the weak bracket condition holds at a point p that is F-accessible from M^* and that condition **(i)** or **(ii)** of Theorem 3.4 holds. Then Theorem 1.11 in case **(i)** (respectively Theorem 2.1 in Chapter 2 in case **(ii)**) implies that for all i, $\Psi(p,i)$ (resp. $\Psi(e^\star,i)$) is a Doeblin point, which is accessible for the process $\tilde{Z}$ from $\mathcal{X}_+$. Thus by point **(iv)** of Theorem 1.13, for all $z=(\rho,\theta,i)\in\mathcal{X}_+$

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

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

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

Then Theorem 3.4 is proved. □

Proof of Theorem 3.6

It suffices to show that Theorems 1.13 and 1.14 remain valid under Assumptions 3.1 and 3.2. For Theorem 1.13, we show that Assumption 1.4 and condition **(e)** of Assumption 1.5 are satisfied. The only difference between Assumptions 3.2 and 1.4 and is that W_K has to be in $\mathcal{D}^2$. so we are done if we prove that $C_c^1\subset\mathcal{D}^2$, which is equivalent to $C_c^1\subset\mathcal{D}$. This is the case by Proposition 1.9. Furthermore, since $\mathcal{X}_0$ is compact, one can modify V and H outside a neighbourhood of $\mathcal{X}_0$ in such a way that all the conditions of Assumption 1.5 hold, in particular Condition **(e)** is checked (see [Ben18, Remark 19].)

For Theorem 1.14, we note that point (i) and (ii) in its proof are still valid since in our case, the set $\mathcal{X}_0$ is compact (see Propositions 8.2 and 8.3 in [Ben18]). Thus, point **(i)** of Theorem 1.14 can be shown by the same argument even if $\mathcal{X}$ is not compact. Now, the existence of a Lyapunov function implies that there exists a compact set $K\subset M$ containing 0, such that, for all $(x,i)\in M\times E$, $\mathbb{P}_{(x,i)}(T_K<\infty)=1$, where T_K is the hitting time of K. Moreover, due to the accessibility of 0, for all neighbourhood U of 0, there exists $\delta>0$ such that, for all $(x,i)\in K\times E$, $\mathbb{P}_{(x,i)}(T_U<\infty)\ge\delta$. Hence, by Markov property, $\mathbb{P}_{(x,i)}(T_U<\infty)\ge\delta$ for all $(x,i)\in M\times E$ and point **(ii)** of Theorem 1.14 follows. □

3.6 Proof of Theorem 3.10

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 [Chu02]. 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, [Chu02, 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 $t\ge 0$, for all $\omega\in\Omega$, for all $D\subset\{1,\dots,d\}$, for all $x,y\in\mathbb{R}^d_{++,D}$,

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

and the inequality is strict if $D = \{1, \ldots, d\}$, $x \neq y$ 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 3.8. *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

$$\sup_{x,y \in \Delta_\varepsilon(K)} \frac{p(\varphi(x), \varphi(y))}{p(x,y)} \leq \alpha, \tag{3.25}$$

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 [Chu02, Proposition 4.1.1] or [BS09, Proposition 6]). Componentwise, this means that for all i,

$$\frac{\langle \nabla \varphi_i(x), x \rangle}{\varphi_i(x)} < 1. \tag{3.26}$$

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,

$$|y_k - x_k| \leq x_k(1 + Mp(x,y))p(x,y). \tag{3.27}$$

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 (3.27) 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 (3.26), continuity of φ and compactness of K imply 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,

$$\sup_{x,y \in \Delta_\varepsilon^c(K)} \frac{p(\varphi(x), \varphi(y))}{p(x,y)} \leq \tau'. \tag{3.28}$$

Combining (3.25) and (3.28) gives the result with $\tau_K(\varphi) = \max(\alpha, \tau') < 1$. □

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 3.10 is a consequence of the weak form of Harris' theorem due to Hairer, Mattingly and Scheutzow [HMS11, Theorem 4.8 and remark 4.10]. More precisely, it states that point **(i)** of Theorem 3.10 holds, provided the three following assumptions are verified (here we let P_t denoted P_t^Z) :

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 3.9. In [HMS11, 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 [HMS11] 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 3.10 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 3.10

A1 We have the following lemma :

Lemma 3.9. *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 [BL16, Lemma 3.5]. □

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 [AC12, 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 $t \in [t_*, t^*]$,

$$\sup_{x,y \in M_v; \omega, \omega' \in \Omega} p(\Psi(t, \omega, x), \Psi(t, \omega', y)) \leq \sup_{a,b \in \mathbf{\Psi}_{[t_*,t^*]}(M_V)} p(a, b) = K. \tag{3.29}$$

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 [BLBMZ12]. As stated in [BLBMZ12, 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 (3.29). 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 $p(\Psi(t, \omega, x), \Psi(t, \omega, y)) \leq p(x, y)$, 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

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

where $\tau_{\mathbf{\Psi}_{t_0}(M_V)}(\Psi(t-t_0,\Theta_{t_0}\omega)) < 1$ is the contraction constant given by Lemma 3.8 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

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

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)**. □

3.7 Appendix

3.7.1 Proof of Proposition 3.3

Recall (see section 3.4) 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 [Sen06]) defined by

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

Note that

$$d_H(\frac{x}{\|x\|},\frac{y}{\|y\|}) = d_H(x,y) \tag{3.30}$$

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

$$\|x-y\| \leq \mathrm{e}^{d_H(x,y)} - 1. \tag{3.31}$$

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 $x, y \in \mathbb{R}^d_{++}$

$$d_H(Tx,Ty) \leq \tau[T] d_H(x,y) \tag{3.32}$$

The number $\tau[T]$ is usually called the *Birkhoff's contraction coefficient* of T, and is given by an explicit formulae (see e.g [Sen06], 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 (3.32) extends to $\mathcal{M}_+$.

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

$$\forall t \geq 0,\ \frac{dM}{dt} = A^{\omega_t} M, M_0 = I_d. \tag{3.33}$$

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 $\mathbb{P}_i^J(\tilde{\Omega}) = 1$ 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 $x, y \in \mathbb{R}^d_+ \setminus \{0\}$

$$\limsup_{t\to\infty} \frac{1}{t} \log d_H(\varphi(t+n,\omega)x, \varphi(t+n,\omega)y) < 0. \tag{3.34}$$

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 (3.34), (3.30), (3.31) 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,

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

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 μ, $\nu \in \mathcal{P}_{inv}^{(\Theta,J)}$. Then one has

$$\lim_{t\to\infty} \sum_i p_i \int_{(S^{d-1}_+)^2} \left(P_t^{(\Theta,J)} f(x,i) - P_t^{(\Theta,J)} f(y,i)\right) \mu(\mathrm{d}x|i)\nu(\mathrm{d}y|i) = 0, \tag{3.35}$$

where $\mu(\cdot|i) = \mu^i(\cdot)/p_i$. But by invariance of μ and ν, the left-hand side of (3.35) 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 (see Theorem 1.8), applied to the PDMP (3.33), 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, $\varphi(1, \boldsymbol{\Theta}_n(\omega)) \in \mathcal{U}$ for infinitely many n, and consequently, for n large enough

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

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

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

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

Replacing ω par $\boldsymbol{\Theta}_{\mathbf{n}}(\omega)$ proves assertion (ii). □

3.7.2 Proof of Lemma 3.3

Before proving Lemma 3.3, we prove the following lemma, which is a consequence of results from Freidlin and Wentzell [FW12].

Lemma 3.10. *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 (3.2). 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 $T > 0$,*

$$\lim_{\varepsilon\to 0} \mathbb{P}_{(x,i)}\left(\max_{0\le t\le T}|X_t^\varepsilon - \Psi_t(x)| > \delta\right) = 0, \tag{3.36}$$

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

Proof According to [FW12, Chapter 2 Theorem 1.3], it suffices to show that for all $\delta > 0$ and all $T > 0$,

$$\lim_{\varepsilon\to 0} \mathbb{P}_i^J\left(\left|\int_{t_0}^{t_0+T}(F^{J_t^\varepsilon}(x) - F^p(x))\mathrm{d}t\right| > \delta\right) = 0, \tag{3.37}$$

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

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

so (3.37) 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_jT uniformly in $t_0 > 0$. By Fubini's Theorem and invariance of p, $\mathbb{E}_p^J\left(\int_{t_0}^{t_0+T}\mathbb{1}_{J_t^\varepsilon=j}\mathrm{d}t\right) = p_jT$, so Bienaymé - Tschebischev inequality gives

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

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 $(p_jT)^2$ 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. $\square$

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 [Ben98, Corollary 3.2].

Lemma 3.11. *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 $\varepsilon>0$,

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

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 3.10. $\square$

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 3.3. For every continuous

$f : M \times E \to \mathbb{R}$, every $t \geq 0$ 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

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

where the right hand side converges to 0 when ε goes to 0 thanks to Lemma 3.10, so A converges to 0. Next,

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

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. □

Acknowledgments

This work was supported by the SNF grant 2000020 − 149871/1. We thank Janusz Mierczyński for many valuable comments on section 2, especially proof of proposition 3.1. ES thanks Carl-Erik Gauthier for many discussions on this subject.

Chapter 4

Application to a predator-prey system with random switching

We reproduce here the article [HS19] which is a joint work with Alexandru Hening, accepted for publication at *SIAM Journal on Mathematical Analysis*. In this chapter, we use some results of Chapter 3 to prove a conjecture by Takeuchi et al. [TDHS06] on a predator-prey Lotka-Volterra system in a random environment. More precisely, we consider the system obtained by a random switching between two vector fields of the form

$$F^i(x,y) = \begin{pmatrix} x(a_i - b_i y) \\ y(-c_i + d_i x) \end{pmatrix}.$$

In the case when the equilibrium points of the two deterministic Lotka-Volterra systems coincide we show that almost surely the trajectory does not converge to the common deterministic equilibrium. Instead, with probability one, the densities of the prey and the predator oscillate between 0 and ∞.

The proof of the conjecture is a corollary of a result we prove about linear switched systems. Assume (Y_t, I_t) is a PDMP that evolves according to $\frac{dY_t}{dt} = A_{I_t} Y_t$ where A_0, A_1 are 2×2 matrices and I_t is a Markov chain on $\{0,1\}$ with transition rates $k_0, k_1 > 0$. If the matrices A_0 and A_1 are not proportional and are of the form

$$A_i := \begin{pmatrix} a_i & b_i \\ c_i & -a_i \end{pmatrix},$$

with $a_i^2 + b_i c_i < 0$, then the average growth rates of $\|Y_t\|$ are all equal and strictly positive. In particular, almost surely $\lim_{t\to\infty} \|Y_t\| = +\infty$.

Keywords: Piecewise deterministic Markov processes; random switching; Lyapunov Exponents; population dynamics; Lotka-Volterra; telegraph noise

MSC primary: 60J99, 34F05, 37H15, 37A50, 92D25

Contents

4.1 Introduction and main results

One of the key issues in ecology is determining when species will persist and when they will go extinct. The randomness of the environment makes the dynamics of populations inherently stochastic and therefore we need to take into account the combined effects of biotic interactions and environmental fluctuations. One way of doing this is by modelling the densities of various species as Markov processes and looking at their long-term behavior (see [Che00, ERSS13, EHS15, LES03, SLS09, SBA11, BEM07, BS09, BHS08, CM10, HNY18a, HN]).

In order to allow for environmental fluctuations and their effect on the persistence or extinction of species one approach is to study stochastic differential equations ([ERSS13, SBA11, HNY18a, HN, HN18b, HN18a]). The other possible approach is to look at stochastic equations driven by a Markov chain. These systems are sometimes called *Piecewise Deterministic Markov Processes (PDMP)* or systems with *telegraph noise*.

For a predator-prey system the classical deterministic example is the Lotka-Volterra model (see [Lot25] and [Vol28])

$$\begin{aligned} \frac{dx(t)}{dt} &= x(t)(a - by(t)), \\ \frac{dy(t)}{dt} &= y(t)(-c + dx(t)), \end{aligned} \tag{4.1}$$

where $x(t), y(t)$ are the densities of the prey and the predator at time $t \geq 0$ and a, b, c and d are positive constants. If one assumes that $x(0) = x_0 > 0, y(0) = y_0 > 0$, so that both predator and prey are present, then the solutions of system (4.1) are periodic (see [Gil75, HS98]) and given in phase space by the curves described by the first integral,

$$r(x, y) = dx - c - c\ln(1 + (dx - c)/c) + by - a - a\ln(1 + (by - a)/a) = \text{constant} = r. \tag{4.2}$$

One should note that both the predator and the prey from (4.1) do not experience intraspecific competition. In particular, if the predator is not present (i.e. $y_0 = 0$) then the prey density blows up to infinity. In [GH79, MHP14] the authors are able to analyze the n-dimensional generalization of (4.1) i.e. the setting when one has one prey and $n - 1$ predators and each species interacts only with the adjacent trophic levels. Stochastic predator-prey models have been studied in the stochastic differential equation setting by [Rud03, HN18b, HN18a]. However, we note that in all these studies one needed to assume that there exists intraspecific competition among the prey and the predators. This simplifies the analysis significantly because the predator and the prey densities get pushed towards the origin when they become too large.

In [AHS79] the authors show that if the coefficient a (growth rate of the prey) is randomly perturbed by white noise then the resulting stochastic system cannot have a stationary distribution and that as the time goes to infinity, with probability 1, explosion does not occur. In [KK01] the authors look at scaling limits of Lotka-Volterra systems perturbed by white noise - they prove that a suitably rescaled version of $r(x(t), y(t))$, where $r(x, y)$ is the first integral from (4.2), converges to a one-dimensional stochastic differential equation. They then use this SDE to gain information about both the deterministic and the stochastic Lotka-Volterra systems.

Like in [TDHS06], we consider the random switching between two Lotka-Volterra prey-predator systems of the form (4.1). More precisely, for $i \in E := \{0, 1\}$, let $F^i : \mathbb{R}^2_+ \to \mathbb{R}^2_+$

denote the vector field

$$F^i(x,y) = \begin{pmatrix} x(a_i - b_i y) \\ y(-c_i + d_i x) \end{pmatrix} \tag{4.3}$$

with $a_i, b_i, c_i, d_i > 0$. Let $(I_t)_{t\geq 0}$ be a continuous-time Markov Chain defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $E := \{0,1\}$. Suppose I_t has transition rates $k_0, k_1 > 0$. Throughout the paper we will let $\mathbb{R}^2_{++} := \{(x_1,x_2) \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0\}$ and $\mathbb{R}^2_{+} := \{(x_1,x_2) \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0\}$. We denote by $(X_t)_{t\geq 0} = (x_t, y_t)_{t\geq 0}$ the solution of

$$\begin{aligned} \frac{dx_t}{dt} &= x_t(a_{I_t} - b_{I_t} y_t) \\ \frac{dy_t}{dt} &= y_t(-c_{I_t} + d_{I_t} x_t) \end{aligned} \tag{4.4}$$

for some initial condition $X_0 = (x_0, y_0) \in \mathbb{R}^2_{++}$. The process $(X,I) = (X_t, I_t)_{t\geq 0}$ is a Piecewise Deterministic Markov Process as introduced in Section 1.2.4.

As usual, for $\mathbf{x} \in \mathbb{R}^2$ and $i \in E$, we denote by $\mathbb{P}_{\mathbf{x},i}$ the law of the process (X,I) when $(X_0, I_0) = (\mathbf{x}, i)$ almost surely and by $\mathbb{E}_{\mathbf{x},i}$ the associated expectation.

The vector field F^i from (4.3) has a unique positive equilibrium $(p_i, q_i) = (c_i/d_i, a_i/b_i)$. In [TDHS06] the authors look at the two cases

I. $p_0 = p_1 =: p$ and $q_0 = q_1 =: q$, i.e. common zero for F^0 and F^1,

II. $(p_0, q_0) \neq (p_1, q_1)$, i.e. different zeroes for F^0 and F^1.

We assume throughout this paper that $p_0 = p_1 =: p$ and $q_0 = q_1 =: q$. The vector fields F^0 and F^1 therefore have a common zero - this will allow us to use the results from Chapter 3. We also assume that F^0 and F^1 are non collinear to avoid trivial switching.

In [TDHS06, Theorem 4.5] it is shown that only two long term behaviours are possible when the vector fields have a common zero: either X_t converges almost surely to the common equilibrium (p,q), or each coordinate oscillates between 0 and $+\infty$.

Theorem 4.1 (Takeuchi et al., 2006)**.** *For any $(x_0, y_0) \in \mathbb{R}^2_{++}$, with probability 1, either*

$$\lim_{t\to\infty} X_t = (p,q), \tag{4.5}$$

or

$$\limsup x_t = \limsup y_t = +\infty, \quad \liminf x_t = \liminf y_t = 0. \tag{4.6}$$

It was conjectured from simulations (see [TDHS06, Remark 5.1]) that only case 4.6 happens in the above theorem. Using Theorem 4.4 below and results from Chapter 3, we are able to prove this conjecture:

Theorem 4.2. *There exist $\varepsilon > 0$, $b > 1$, $\theta > 0$ and $c > 0$ such that for all $\mathbf{x} := (x_0, y_0) \in \mathbb{R}^2_{++} \setminus \{(p,q)\}$ and $i \in E$,*

$$\mathbb{E}_{\mathbf{x},i}(b^{\tau^\varepsilon}) \leq c(1 + \|\mathbf{x} - (p,q)\|^{-\theta}),$$

where $\tau^\varepsilon := \inf\{t \geq 0 : \|X_t - (p,q)\| \geq \varepsilon\}$. In particular, for any $(x_0, y_0) \in \mathbb{R}^2_{++} \setminus \{(p,q)\}$ we have with probability 1 that

$$\limsup_{t\to\infty} x_t = \limsup_{t\to\infty} y_t = +\infty, \quad \liminf_{t\to\infty} x_t = \liminf_{t\to\infty} y_t = 0.$$

Our result provides a deeper understanding of Lotka-Volterra systems in random environments, continuing the work started in [KK01] and [AHS79].

Actually, thanks to Theorem 4.4, the first part of Theorem 4.2 can be generalised as follows. For $i \in E$, let F^i be a vector field of class C^2 on $\mathbb{R}^2$, such that $F^i(0) = 0$. Also assume that for $i \in E$, $DF^i(0)$, the Jacobian matrix of F at 0, has two purely imaginary eigenvalues. In this case, the equilibrium 0 is sometimes called a *center*. We now consider a Markov process $(X_t, I_t)_{t\geq 0}$ where X_t is solution of

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

and I_t is a jump process on E whose rates depend on X

$$\mathbb{P}(I_{t+s} = 1-i | I_t = i, \mathcal{F}_t) = k_{i,1-i}(X_t)s + o(s),$$

where $F_t = \sigma((X_s, I_s) : \ s \leq t)$ and for all x, $(k_{ij}(x))_{i,j}$ is an irreducible matrix that is continuous in x. The process (X, I) is still a PDMP. We can prove the following (see Remark 4.2) in this more general setting.

Theorem 4.3. *Assume that* 0 *is a center for* F^0 *and* F^1 *and that* $DF^0(0)$ *and* $DF^1(0)$ *are non proportional. Then there exist* $\varepsilon > 0$, $b > 1$, $\theta > 0$ *and* $c > 0$ *such that for all* $\mathbf{x} := (x_0, y_0) \in \mathbb{R}^2 \setminus \{(0,0)\}$ *and* $i \in E$,

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

where $\tau^\varepsilon = \inf\{t \geq 0 : \ \|X_t\| \geq \varepsilon\}$. *In particular, for any* $(x_0, y_0) \in \mathbb{R}^2 \setminus \{(0,0)\}$, *with probability one,* X_t *cannot converge to* $(0,0)$.

In view of Theorem 3.5, a strategy to prove the above theorem is to show that the extremal growth rate is positive (see Section 3.2 in Chapter 3.) This is what we do in the next section.

4.2 A result on linear switched systems

Let A_i denote the matrix

$$A_i := \begin{pmatrix} a_i & b_i \\ c_i & -a_i \end{pmatrix}, \tag{4.7}$$

for $i = 0, 1$, where a_i, b_i, c_i are real numbers satisfying

$$a_i^2 + b_i c_i < 0. \tag{4.8}$$

In this case, both matrices A_0, A_1 have purely imaginary eigenvalues.

We want to investigate the sign of the average growth rates associated to A_0 and A_1 under the switching generated by $(I_t)_{t\geq 0}$. Recall (see Section 3.2 in Chapter 3) that these quantities are defined by

$$\Lambda(\mu) := \int \langle A_i\theta, \theta\rangle \mu(d\theta di),$$

where μ is an invariant probability measure of the process $(\Theta_t, I_t)_{t\geq 0}$ on $S^{d-1} \times E$. Here Θ_t stands for the solution of

$$\frac{d\Theta_t}{dt} = A_{I_t}\Theta_t - \langle A_{I_t}\Theta_t, \Theta_t\rangle\Theta_t.$$

The *extremal average growth rates* are the numbers defined by

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

where $\mathcal{P}_{erg}$ is the set of ergodic measures of $((\Theta_t, I_t))_{t\geq 0}$ on $S^1 \times E$.

The main result of this section is the following theorem.

Theorem 4.4. *Assume A_0 and A_1 are non proportional matrices of the form* (4.7) *with coefficients satisfying* (4.8). *Then all the average growth rates are equal and strictly positive*

$$\Lambda^+ = \Lambda^- > 0.$$

As shown in the following lemma, non proportionality is not required to prove the uniqueness and nonnegativity of the average growth rate.

Lemma 4.1. *The process (Θ_t, I_t) admits a unique invariant probability measure μ on $S^1 \times E$. Furthermore, $\Lambda(\mu) \geq 0$.*

Proof. The uniqueness follows from Proposition 3.2 and Example 3.2. Indeed, since we study a two dimensional system, a sufficient condition is that at least one matrix A_i has no real eigenvalue. This is the case for both A_1 and A_2. Since A_0 and A_1 have zero trace, Corollary 3.1 implies that $\Lambda(\mu) = \Lambda^+ \geq 0$. $\square$

4.2.1 Proof of Theorem 4.4

In order to prove Theorem 4.4, we will use Theorem 1.19 due to Bougerol [Bou88] on Lyapunov exponents. Recall that according to Proposition 1.15, there exist $d \in \{1,2\}$ numbers $\lambda_1 > \lambda_d$ called *the Lyapunov exponents*, a Borel set $\tilde{\Omega} \subset \Omega$ with $\mathbb{P}(\tilde{\Omega}) = 1$, and for each $\omega \in \tilde{\Omega}$ distinct vector spaces

$$\{0\} = V_{d+1}(\omega) \subset V_d(\omega) \subset V_1(\omega) = \mathbb{R}^2$$

such that

$$\lim_{t\to\infty} \frac{1}{t} \log \|Y_t\| = \lambda_i \tag{4.9}$$

for all $\mathbf{y}_0 \in V_i(\omega) \setminus V_{i+1}(\omega)$, where Y_t is solution to

$$\begin{aligned} \frac{dY_t}{dt} &= A_{I_t} Y_t \\ Y_0 &= \mathbf{y}_0 \in \mathbb{R}^2 \setminus \{(0,0)\}. \end{aligned} \tag{4.10}$$

Remark 4.1. *By Proposition 3.1 and Corollary 3.1, one can note that $\sum_i \lambda_i = 0$ and $\lambda_1 = \Lambda^+$. Therefore, proving that $\Lambda^+ > 0$ is equivalent to showing that $d = 2$.*

In order to prove that $d = 2$, we will use [Bou88, Theorem 1.7] that we have already stated in Chapter 1 (see Theorem 1.19).

We show that we can use Theorem 1.19 in our context. Let $(M_t)_{t\geq 0}$ with $M_t \in GL_2(\mathbb{R}), t \geq 0$ be the solution of the matrix equation

$$\begin{aligned} \frac{dM_t}{dt} &= A_{I_t} M_t \\ M_0 &\in GL_2(\mathbb{R}). \end{aligned} \tag{4.11}$$

The process (M_t, I_t) is a PDMP living on $GL_2(\mathbb{R}) \times E$. One can note that when $M_0 = \mathrm{Id}$, the identity matrix, then for all $\mathbf{y} \in \mathbb{R}_+^2$ the process Y from (4.10) can be written as

$$Y_t = M_t \mathbf{y}$$

if $Y_0 = \mathbf{y}$. Recall that the definition of a multiplicative system and hypothesis **H** have been defined in Section 1.4.3 of Chapter 1.

Lemma 4.2. *Set $\pi = (k_1/(k_0+k_1), k_0/(k_0+k_1))$ and $\mathcal{E} = E = \{0,1\}$. Then (M, I, π) is a multiplicative system satisfying **H**.*

Proof. First we show that (M, I, π) is a multiplicative system. (M, I) is a PDMP thus a Markov process. In addition, if we denote by M^N the process M when $M_0 = N$ almost surely, one can easily check that $M^N = M^{\mathrm{Id}} N$ almost surely. As a result we see that point **(ii)** of definition 1.20 is satisfied. Straightforward computations show that π is the unique invariant distribution of I and is therefore ergodic. Let K be a constant such that $\|A_i\| \leq K$ for $i \in E$. Then from

$$\begin{aligned} \frac{dM_t}{dt} &= A_{I_t} M_t, \\ \frac{dM_t^{-1}}{dt} &= -M_t^{-1} A_{I_t} \end{aligned}$$

together with $M_0 = M_0^{-1} = Id$ and Gronwall's Lemma, one can show that for all $t \geq 0$

$$\|M_t\|, \|M_t^{-1}\| \leq e^{Kt}.$$

This proves point **(iii)**.

Now we show that (M, I, π) satisfy hypothesis **H**. In our case, $\mathcal{E} = E$ is a finite set, thus points (i) and (iii) of Definition 1.21 are straightforward. Furthermore, (M, I) is Feller, by Proposition 1.9. $\square$

Thus, Theorem 1.19 apply. For convenience, we state here the result adapted to our context. We denote by U the first order resolvent of $(P_t)_{t\geq 0}$, which is defined via

$$U = \int_0^{+\infty} \mathrm{e}^{-t} P_t \mathrm{d}t,$$

where $(P_t)_{t\geq 0}$ is the semigroup of the PDMP $(M_t, I_t)_{t\geq 0}$ on $GL_2(\mathbb{R}) \times E$. For $i \in E$ let D_i be the support of $U((\mathrm{Id}, i), \cdot)$ and $S_i = \{A \in GL_2(\mathbb{R}) \,:\, (A, i) \in D_i\}$.

Theorem 4.5 (Bougerol, 1988)**.** *Assume (M, σ, π) is a multiplicative system satisfying hypothesis **H**. Assume furthermore that*

(i) *There exists a matrix in S_1 with two eigenvalues with different modulus.*

(ii) *There does not exist some finite union W of one-dimensional vector spaces such that, for all matrices M in S_1, $MW = W$.*

Then $d = 2$.

We now pass to the proof of Theorem 4.4. We start by showing that it suffices to prove Theorem 4.4 for a specific class of matrices.

Lemma 4.3. *Assume Theorem 4.4 holds when A_0 is of the special form*

$$A_0 = \begin{pmatrix} 0 & -\omega_0 \\ \omega_0 & 0 \end{pmatrix}.$$

Then Theorem 4.4 holds for any A_0.

Proof. First we show that a linear change of coordinates does not change the value of Λ^+. Let $G \in GL_2(\mathbb{R})$, and set, for all $t \geq 0$, $Z_t = GY_t$. Then (Z_t, I_t) is a PDMP with Z solution of

$$\frac{dZ_t}{dt} = B_{I_t} Z_t,$$

where $B_i = GA_iG^{-1}$. Due to $\lambda_1 = \frac{\log \|Y_t\|}{t} = \lim \frac{\log \|Z_t\|}{t}$ one can see that the maximal growth rates of Y_t and Z_t are equal.

Next, since the eigenvalues of A_0 are $\pm i\omega_0$ for $\omega_0 := \sqrt{-(a_0^2 + b_0c_0)}$, a classical result in linear algebra (see for example [HS74, Chapter 4, Theorem 3]) states that there exists a matrix $G \in GL_2(\mathbb{R})$ such that

$$B_0 = GA_0G^{-1} = \begin{pmatrix} 0 & -\omega_0 \\ \omega_0 & 0 \end{pmatrix}.$$

Thus if the result is shown for a matrix A_0 of this form, it will be proven for every matrix A_0 with purely complex eigenvalues because A_0 and A_1 are proportional if and only if B_0 and B_1 are. □

Proof of Theorem 4.4. It suffices to show that **(i)** and **(ii)** of Theorem 4.5 are satisfied.

We first show that **(i)** holds.

According to Lemma 4.3, it suffices to show the assumptions are satisfied for A_0 of the form

$$A_0 = \begin{pmatrix} 0 & -\omega_0 \\ \omega_0 & 0 \end{pmatrix}.$$

By standard computations, one can show that for all $t \geq 0$ and $i \in E$,

$$e^{tA_i} = \cos(\omega_i t)\mathrm{Id} + \frac{1}{\omega_i}\sin(\omega_i t)A_i,$$

where $\omega_i := \sqrt{-(a_i^2 + b_ic_i)}$. In particular, since $\mathrm{Tr}(A_i) = 0$, one has that for all $s, t \geq 0$

$$\varphi(s,t) := \mathrm{Tr}(e^{sA_0}e^{tA_1}) = 2\cos(\omega_0 s)\cos(\omega_1 t) + \frac{1}{\omega_0\omega_1}\sin(\omega_0 s)\sin(\omega_1 t)\mathrm{Tr}(A_0A_1).$$

On the other hand, since $\mathrm{Tr}(A_i) = 0$, one has $\det(e^{sA_0}e^{tA_1}) = 1$. Thus, denoting by μ_1, μ_2 the eigenvalues of $e^{sA_0}e^{tA_1}$ one can see that $\mu_1\mu_2 = 1$ or equivalently $\mu_1 = 1/\mu_2$. In order to apply Theorem 4.5, we need to have $|\mu_1| > |\mu_2|$. This cannot happen if $\mu_1 = \mu_2$ or if $\mu_1 = \bar{\mu}_2$.

Due to the fact that $\mu_1 + \mu_2 = \mathrm{Tr}(e^{sA_0}e^{tA_1})$, the condition $|\mu_1| > |\mu_2|$ is equivalent to $|\varphi(s,t)| > 2$. By studying the derivatives of $\varphi(s,t)$, one sees that its extremal values are reached at points (s^*, t^*) of the form $(\pi/\omega_0, \pi/\omega_1)$ or $(\pi/2\omega_0, \pi/2\omega_1)$ modulo π. From this we note that the extremal values are $\varphi(s^*, t^*) = \pm 2$ and

$$\varphi(s^*, t^*)^2 = \frac{1}{\omega_0^2\omega_1^2}\mathrm{Tr}(A_0A_1)^2 = \frac{(b_1 - c_1)^2}{\omega_1^2}.$$

Therefore

$$\varphi(s^*, t^*)^2 > 4 \iff a_1 > 0 \text{ or } b_1 + c_1 > 0 \iff A_1 \text{ is not proportional to } A_0. \tag{4.12}$$

By assumption, A_1 and A_0 are not proportional. Therefore, using (4.12) one infers that the matrix $N(t_0, t_1) = N := e^{t_0A_0}e^{t_1A_1}$ has two eigenvalues with different moduli for $(t_0, t_1) = (\pi/2\omega_0, \pi/2\omega_1)$.

In order to conclude that assumption **(i)** from Theorem 4.5 is satisfied, we show that the matrix N lies in S_1.

Let V be a neighborhood of N in $GL_2(\mathbb{R})$. Then, by continuity, there exists $\varepsilon > 0$ such that for all $u \in [t_0 - \varepsilon, t_0 + \varepsilon]$, $s \in [t_1 - \varepsilon, t_1 + \varepsilon]$ and $\delta \leq \varepsilon$, the matrix $N_{s,u,\delta} = e^{\delta A_1}e^{uA_0}e^{sA_1}$ is in V. Let V_ε be the set of the matrices $N_{s,u,\delta}$ for s, u and δ as before. Let $(U_n)_{n\geq 1}$ denote the sequence of interjump times of the process I. Then, on the event $B_{t,\varepsilon} = \{U_1 \in [t_0 - \varepsilon, t_0 + \varepsilon]; U_2 \in [t_1 - \varepsilon, t_1 + \varepsilon]; t - (U_1 + U_2) \leq \varepsilon; U_1 + U_2 + U_3 \geq t\}$, $I_t = 1$ and $M_t \in V_\varepsilon$. Thus one has

$$\begin{aligned}\mathbb{P}_{\mathrm{Id},1}\left((M_t, I_t) \in V \times \{1\}\right) &\geq \mathbb{P}_{\mathrm{Id},1}\left((M_t, I_t) \in V_\varepsilon \times \{1\}\right)\\ &\geq \mathbb{P}_{\mathrm{Id},1}\left(B_{t,\varepsilon}\right).\end{aligned}$$

This last probability is positive for all $\varepsilon > 0$ and $t \in [t_0 + t_1 - 2\varepsilon, t_0 + t_1 + 3\varepsilon]$. Hence

$$U((\mathrm{Id}, 1), V \times \{1\}) = \int_0^{+\infty} \mathrm{e}^{-t}\mathbb{P}_{\mathrm{Id},1}\left((M_t, I_t) \in V \times \{1\}\right) dt > 0.$$

This is true for all neighborhoods of N, so $N \in S_1$ and point **(i)** is shown.

Using similar arguments, one can show that the family of matrices $\left(e^{tA_1}\right)_{t\geq 0}$ is in S_1. Since A_1 has two complex eigenvalues, one cannot find a finite union of one dimensional vector spaces invariant by the family $\left(e^{tA_1}\right)_{t\geq 0}$. This proves that assumption **(ii)** of Theorem 4.5 holds. □

4.3 Proof of Theorem 4.2

Let A_i denote the Jacobian matrix of the vector field F^i at (p, q). Then

$$A_i = \begin{pmatrix} 0 & -b_ip \\ d_iq & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\alpha_i \\ \beta_i & 0 \end{pmatrix},$$

where $\alpha_i = b_i p$ and $\beta_i = d_i q$. The linear PDMP (Y, I) where Y is the solution of

$$\frac{dY_t}{dt} = A_{I_t} Y_t,$$

is a particular case of the systems studied in Section 4.2.

To apply Theorem 4.4, we have to check that A_0 and A_1 are non collinear. This is equivalent to showing that $\alpha_1\beta_0 \neq \alpha_0\beta_1$. Assume that $\alpha_1\beta_0 = \alpha_0\beta_1$. Then since $\alpha_i = b_i p$ and $\beta_i = d_i q$, we get $b_1 d_0 = b_0 d_1$. Moreover, since $p_0 = p_1$ and $q_0 = q_1$, one has $c_0 d_1 = c_1 d_0$ and $a_0 b_1 = a_1 b_0$. If we set $\gamma = b_1/b_0$, we note that $k\kappa_1 = \gamma\kappa_0$ for $\kappa = a, b, c, d$, which implies $F^1 = \gamma F^0$. This contradicts the assumption that the vector fields F^0 and F^1 are non collinear.

As a result, A_0 and A_1 cannot be collinear. We can therefore apply Theorem 4.4 and conclude that $\Lambda^- = \Lambda^+ > 0$.

We now prove rigorously point **(ii)** of Theorem 3.5 of Chapter 3 in the present context.

Let $K \subset \mathbb{R}^2_{++}$ be a compact set containing (p, q) in its interior. Let $\varphi_K : \mathbb{R}^2_{++} \to [0, 1]$ be a smooth function such that $\varphi_K = 1$ on K^δ and $\varphi_K = 0$ on the complement of $K^{2\delta}$. Here $K^\delta = \{x \in \mathbb{R}^2_{++} \ : \ d(x, K) < \delta\}$ is the δ - neighbourhood of K and $\delta > 0$ is such that $K^{2\delta} \subset \mathbb{R}^2_{++}$. For $i \in E$, set $F^i_K = \varphi_K F^i$. Note that $F^i_K = F^i$ on K^δ. In particular, (p, q) is a common zero of F^0_K and F^1_K and $DF^i_K((p, q)) = DF^i((p, q)) = A_i$. Now consider the PDMP (X^K, I), with $(X^K_t)_{t\geq 0}$ solution of

$$\frac{dX^K_t}{dt} = F^{I_t}_K(X^K_t).$$

Then we have the two following facts. First, denote by $\tau_K = \inf\{t \geq 0 : \ X_t \notin K\}$ the exit time of K for X_t. Then if $X_0 = X^K_0 = \mathbf{x} \in K$, for all $t \leq \tau_K$, $X_t = X^K_t$ almost surely. Next, since $DF^i_K((p, q)) = A_i$, the average growth rate Λ^-_K of (X^K, I) is equal to Λ^-. Now, since X^K_t remains in the compact set $\overline{K^{2\delta}}$ and $\Lambda^-_K = \Lambda^- > 0$, one can apply Theorem 3.2. According to point **(iii)** of this theorem, since $\Lambda^-_K > 0$, there exist $\varepsilon > 0$, $\theta > 0$, $b > 1$ and $c > 0$ such that for all $\mathbf{x} \in \mathbb{R}^2_{++} \setminus \{(p, q)\}$ and $i \in E$,

$$\mathbb{E}_{\mathbf{x},i}(b^{\tau^\varepsilon_K}) \leq c(1 + \|\mathbf{x} - (p, q)\|^{-\theta}),$$

where $\tau^\varepsilon_K = \inf\{t \geq 0 : \|X^K_t - (p, q)\| \geq \varepsilon\}$. Without loss of generality, we can assume that the ball of center (p, q) and radius ε is included in the interior of K. Let $\tau^\varepsilon = \inf\{t \geq 0 : \|X_t - (p, q)\| \geq \varepsilon\}$. Now if $\|\mathbf{x} - (p, q)\| \geq \varepsilon$, $\tau^\varepsilon = 0$. If $\|\mathbf{x} - (p, q)\| < \varepsilon$, then since $X_t = X^K_t$ for all $t \leq \tau_K$, one gets that $\tau^\varepsilon = \tau^\varepsilon_K \leq \tau_K$. In particular, for all $\mathbf{x} \in \mathbb{R}^2_{++} \setminus \{(p, q)\}$ and $i \in E$,

$$\mathbb{E}_{\mathbf{x},i}(b^{\tau^\varepsilon}) \leq c(1 + \|\mathbf{x} - (p, q)\|^{-\theta}). \tag{4.13}$$

We claim that because of (4.13) X_t cannot converge to (p, q). We argue by contradiction. Let $\mathbf{x} \in \mathbb{R}^2_{++}, i \in E$ and assume that X_t converges to (p, q) almost surely under $\mathbb{P}_{\mathbf{x},i}$. Define two stopping times by

$$\tau^{\text{in},1}_{\varepsilon/2} = \inf\{t \geq 0 \ : \ \|X_t - (p, q)\| \leq \varepsilon/2\}$$

and

$$\tau^{\text{out},1}_{\varepsilon} = \inf\{t > \tau^{\text{in},1}_{\varepsilon/2} \ : \ \|X_t - (p, q)\| \geq \varepsilon\}.$$

Since X_t converges to (p, q) almost surely, one has $\mathbb{P}_{\mathbf{x},i}(\tau^{\text{in},1}_{\varepsilon/2} < \infty) = 1$. Using the strong Markov property at $\tau^{\text{in},1}_{\varepsilon/2}$, one gets

$$\mathbb{P}_{\mathbf{x},i}(\tau^{\text{out},1}_{\varepsilon} < \infty) = \mathbb{E}_{\mathbf{x},i}(\mathbb{P}_{X_{\tau^{\text{in},1}_{\varepsilon/2}}}(\tau^\varepsilon < \infty)) = 1.$$

Construct recursively a family of stopping times

$$\tau_{\varepsilon/2}^{\text{in},k} = \inf\{t > \tau_{\varepsilon}^{\text{out},k-1} \ : \ \|X_t - (p,q)\| \leq \varepsilon/2\}$$

and

$$\tau_{\varepsilon}^{\text{out},k} = \inf\{t > \tau_{\varepsilon/2}^{\text{in},k} \ : \ \|X_t - (p,q)\| \geq \varepsilon\},$$

by repeating the above procedure. Then one gets that for all $k \geq 1$, $\tau_{\varepsilon/2}^{\text{in},k}$ and $\tau_{\varepsilon}^{\text{out},k}$ are finite almost surely. This contradicts the fact that X_t converges to (p,q). As a result we have shown that X_t cannot converge to (p,q).

Remark 4.2. *The proof of Theorem 4.2 above extends verbatim to the proof of Theorem 4.3. The fact that the jump rates now depend on the position does not affect the result because when it comes to the linear system in Theorem 4.4, one just has the constants* $k_{ij}(0)$ *as jump rates (see Section 3.2 of Chapter 3 for details).*

Chapter 5

Randomly switched vector fields sharing a zero on a common invariant face

We consider in this Chapter a Piecewise Deterministic Markov Process given by random switching between finitely many vector fields vanishing at 0. We have shown in Chapter 3 that the behaviour of this process is mainly determined by the signs of average growth rates (Lyapunov exponents). However, results have only been given when all these exponents have the same sign. In this note, we consider the degenerate case where the process leaves invariant some face and results are stated when the Lyapunov exponents are of opposite signs. Applications are given to Lorenz vector fields with switching, and to SIRS model in random environment

This Chapter comes from my article [Str18], which has been submitted for publication.

Keywords: Piecewise deterministic Markov processes, Lyapunov Exponents, Stochastic Persistence, Lorenz vector field, Epidemiology, SIRS

MSC primary:60J25, 34A37, 37H15, 37A50, 92D30

Contents

5.1 Introduction

In this chapter, we consider a Markov process obtained by random switching between finitely many vector fields $F^i : \mathbb{R}^d \to \mathbb{R}^d$, sharing a common equilibrium point q. As we have shown in Chapter 3, the behaviour of the switched system near q is determined by the sign of the average growth rates (see Definition 3.7 and Theorems 3.1, 3.2, 3.3 and 3.4 above).

In the present chapter, we consider the degenerate situation, where the process leaves invariant a face $\{0\} \times \mathbb{R}^m \subset \mathbb{R}^d$ containing q. At first glance, this might seem like a strong assumption. However, when one thinks to ecological or epidemiological models, this assumption is quite natural. For example, if we are looking at an epidemiological model with n groups of susceptible and m groups of infected, with $n + m = d$, then the invariance of the face $\{0\} \times \mathbb{R}^m$ simply means that if there is no infected people at the beginning, there will not be infected people in the future (see Example 5.3.2 below).

Since q belongs to $\{0\} \times \mathbb{R}^m$ which is invariant by F^i, the Jacobian matrices have the form

$$A^i = \begin{pmatrix} B^i & 0 \\ C^i & D^i \end{pmatrix}.$$

We show that if both maximal growth rates associated with B^i and D^i are negative, then the process converges to q; while if all the growth rates associated to B^i are positive and those to D^i are negative, the process admits an invariant probability measure that gives no mass to $\{0\} \times \mathbb{R}^m \subset \mathbb{R}^d$, and hence no mass to q. We also notice that in this last case, the growth rates associated to A^i take positive and negative values, so that the results of Chapter 3 cannot be applied. The chapter is organised as follows. In section 5.2, the main results are stated. The proofs are postponed to section 5.4. In section 5.3, we give several applications. In particular, our result enables us to close a gap in a discussion on random switching between two Lorenz vector fields in [BH12]. We also recover and slightly extend the results on SIRS models with Markov switching given in [LLC17].

5.2 Notations and results

Let $d \geq 1$, $E = \{1, \ldots, N\}$ a finite set and for all $i \in E$, $F^i : \mathbb{R}^d \to \mathbb{R}^d$ a C^2 globally integrable vector field. We denote by φ^i the flow induced by F^i and we assume that there exists a closed set M which is forward invariant for all the vector fields, that is

$$\varphi^i_t(M) \subset M, \quad \forall t \geq 0.$$

For all $x \in M$, we are given an irreducible rate matrix $(a_{ij}(x))_{i,j \in E}$, continuous in x. We consider a Markov process $(Z_t)_{t\geq 0} = (X_t, I_t)_{t \geq 0} \in M \times E$, where X evolves according to

$$\frac{dX_t}{dt} = F^{I_t}(X_t), \tag{5.1}$$

and 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 $\mathcal{F}_t = \sigma((X_s, I_s) \ : s \leq t\})$. The process Z belongs to the class of *Piecewise Deterministic Markov Processes* (PDMP) (see Section 1.2.4 in Chapter 1).

Without loss of generality, we assume that $q = 0$. For n, m such that $n + m = d$, and $x \in \mathbb{R}^d = \mathbb{R}^n \times \mathbb{R}^m$, we set $x = (x_n, x_m)$. The notation 0_k for $k = n, m$ refers to the zero vector of $\mathbb{R}^k$. We also write $F^i(x) = (F^i_n(x), F^i_m(x))$. Our standing assumption is :

Assumption 5.1.

1. *The origin lies in M and for all $i \in E$, $F^i(0) = 0$.*
2. *For all $x_m \in \mathbb{R}^m$ and all $i \in E$, $F^i_n(0_n, x_m) = 0$.*
3. *The set M intersects the face $\{0_n\} \times \mathbb{R}^m$: $\{0\} \subsetneq M \cap (\{0_n\} \times \mathbb{R}^m) \subsetneq M$*
4. *The set M is compact and locally star shaped around the origin : there exists $\delta > 0$ such that*
$$x \in M \text{ and } \|x\| \leq \delta \Rightarrow [0, x] \subset M,$$
where $[0, x] = \{tx,\ t \in [0, 1]\}$.

The second assumption implies that the face $\{0_n\} \times \mathbb{R}^m$ is invariant under each φ^i : for all $t \geq 0$, $\varphi^i_t(x) \in \{0_n\} \times \mathbb{R}^m$ if and only if $x \in \{0_n\} \times \mathbb{R}^m$. We set $M_+ = \{(x_n, x_m) \in M\ :\ x_n \neq 0\}$ and $M_0 = M \setminus M_+$. Both M_0 and M_+ are non empty, and M_0 is invariant for all the flows φ^i. For all $i \in E$, set $A^i = DF^i(0)$, the Jacobian matrix of F^i at 0. The second assumption has also the consequence that A^i is block lower triangular :

$$A^i = \begin{pmatrix} B^i & 0 \\ C^i & D^i \end{pmatrix}, \tag{5.2}$$

with $B^i \in M_n(\mathbb{R})$, $C^i \in M_{m,n}(\mathbb{R})$ and $D^i \in M_m(\mathbb{R})$.

5.2.1 Notation

Throughout this chapter, we will adopt the following notation : $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product in $\mathbb{R}^k$, for $k = n, m, d$; $\| \cdot \|$ the associated norm, and $S^{k-1} = \{x \in \mathbb{R}^k\ :\ \|x\| = 1\}$ the unit sphere.

5.2.2 Linear system and Lyapunov exponents

For a given set of matrices $\hat{A} = (\hat{A}^i)_{i \in E}$ of size $k \times k$, we consider the linear system (Y, J) where Y evolves according to
$$\frac{dY_t}{dt} = \hat{A}^{J_t} Y_t,$$
and J is a continuous time Markov chain on E with transition rate matrix $(a_{ij}(0))_{i,j \in E}$. By irreducibility of $(a_{ij}(0))_{i,j \in E}$, J admits a unique invariant probability measure on E denoted by p.

Whenever the initial condition y_0 is not zero, the angular part of Y_t, $\Theta_t = \frac{Y_t}{\|Y_t\|}$ is well defined, and evolves according to

$$\frac{d\Theta_t}{dt} = \hat{A}^{J_t} \Theta_t - \langle A^{J_t} \Theta_t, \Theta_t \rangle \Theta_t. \tag{5.3}$$

This defines a differential equation on S^{k-1} and the process $(\Theta_t, J_t)_{t \geq 0}$ is a PDMP on $S^{k-1} \times E$. When we need to emphasis the dependence on $(\hat{A}^i)_{i \in E}$, we denote by $\Theta(\hat{A})$

the solution of (5.3). For an invariant probability μ of $(\Theta(\hat{A}), J)$, we define the *μ-average growth rate* as

$$\Lambda_{\hat{A}}(\mu) = \int \langle \hat{A}^i\theta, \theta\rangle \mu(d\theta di) = \sum_{i\in E} \int_{S^{k-1}} \langle \hat{A}^i\theta, \theta\rangle \mu^i(d\theta), \tag{5.4}$$

where $\mu^i(\cdot)$ is the measure on S^{k-1} defined by

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

We let $\Lambda(\hat{A})$ be the set of all the $\Lambda_{\hat{A}}(\mu)$ for μ invariant probability of $(\Theta(\hat{A}), J)$. As in Chapter 3 (see Definition 3.7), we define the *extremal average growth rates* as the numbers

$$\Lambda^-_{\hat{A}} = \inf \Lambda(\hat{A}) \text{ and } \Lambda^+_{\hat{A}} = \sup \Lambda(\hat{A}). \tag{5.5}$$

Recall that we have shown in Proposition 3.1 that $\Lambda(\hat{A})$ is composed of Lyapunov exponents in the sense of Oseledet's Multiplicative Ergodic Theorem (see Theorem 1.17 and Proposition 1.15 in Chapter 1). In particular, $\Lambda(\hat{A})$ is actually a finite set, and the supremum and the infimum in equation (5.5) are maximum and minimum. We start with a lemma.

Lemma 5.1. *Assume that all the A^i have the block triangular form* (5.2)*.Then, with the above notations, $\Lambda(D) \subset \Lambda(A)$.*

Proof Let $\lambda \in \Lambda(D)$ and $\hat{\mu}$ be an invariant probability of $(\Theta(D), J)$ such that $\lambda = \Lambda_D(\hat{\mu})$. For $\Theta \in S^{d-1}$, we write $\Theta = (\Theta^n, \Theta^m)$. We this notation, (5.3) becomes :

$$\begin{cases} \frac{d\Theta^n_t}{dt} = B^{J_t}\Theta^n_t - \left(\langle B^{J_t}\Theta^n_t, \Theta^n_t\rangle + \langle C^{J_t}\Theta^n_t + D^{J_t}\Theta^m_t, \Theta^m_t\rangle\right)\Theta^n_t \\ \frac{d\Theta^m_t}{dt} = C^{J_t}\Theta^n_t + D^{J_t}\Theta^m_t - \left(\langle B^{J_t}\Theta^n_t, \Theta^n_t\rangle + \langle C^{J_t}\Theta^n_t + D^{J_t}\Theta^m_t, \Theta^m_t\rangle\right)\Theta^m_t \end{cases} \tag{5.6}$$

From this equation, one can see that the space $\{(\theta_n, \theta_m) \in S^{d-1} : \theta_n = 0\}$ is invariant, and on that space,$(\Theta(A), J) = (0, \Theta(D), J)$. Now we extend $\hat{\mu}$ to a probability measure μ on $S^{d-1} \times E$ such that $\mu(\{(\theta_n, \theta_m) \in S^{d-1} : \theta_n = 0\} \times E) = 1$ and the marginal of μ on $S^{m-1} \times E$ is $\hat{\mu}$. Then, μ is an invariant probability for $(\Theta(A), J)$, and straightforward computation shows that $\Lambda_A(\mu) = \Lambda_D(\hat{\mu}) = \lambda$. Thus $\lambda \in \Lambda(A)$. $\square$

Remark 5.1. The same proof shows that in case where the A^i are block diagonal, that is $C^i = 0$, then $\Lambda(B) \subset \Lambda(A)$. However, this is not true in general. Here is a counter example in dimension $d = 2$. Let A^i, $i = 0, 1$ be two 2×2 matrices defined by

$$A^i = \begin{pmatrix} b_i & 0 \\ c_i & d_i \end{pmatrix},$$

and assume that $b_i < d_i$ for $i = 0, 1$ as well as $c_0(b_1 - d_1) \neq c_1(b_0 - d_0)$. In particular, $\Lambda^+_B = \sum_i p_i b_i < \sum_i p_i d_i = \Lambda^-_D$. We show that in this case, the set of invariant probability measures of $(\Theta(A), J)$ reduces to $\delta_{(0,1)} \otimes p$ and $\delta_{(0,-1)} \otimes p$; hence $\Lambda(A) = \Lambda(D) \not\supset \Lambda(B)$. Let θ_i be the normalized eigenvector of A^i associated with b_i. Since $c_0(b_1 - d_1) \neq c_1(b_0 - d_0)$, $\theta_0 \neq \theta_1$. Now it is easily checked that the region between θ_0 and θ_1 is transient for $\Theta(A)$ and that when the process leaves this region, $\Theta_t(A)$ converges to $(0, 1)$ or $(0, -1)$.

We prove in Section 5.4 that the result given in the preceding remark can be generalized as follows :

Proposition 5.1. *Assume that all the A^i have the block triangular form* (5.2)*. If $\Lambda_B^+ < \Lambda_D^-$ and if $\{(\theta_n, \theta_m) \in S^{d-1} \,:\, \theta_n = 0\}$ is accessible from S^{d-1}, then $\Lambda(A) = \Lambda(D)$.*

Using Theorem 1.18 due to Hennion [Hen84], we have the following proposition:

Proposition 5.2. *Assume that all the A^i have the block triangular form* (5.2)*.Then, with the above notations, $\Lambda_A^+ = \max(\Lambda_B^+, \Lambda_D^+)$.*

Proof

We let $S(\hat{A})$ denoted the set of Lyapunov exponents counted with multiplicity (see Theorem 1.17) and $\lambda_1(\hat{A}) = \max S(\hat{A})$. By Proposition 3.1, $\Lambda(\hat{A}) \subset \mathcal{S}(\hat{A})$ and $\Lambda_{\hat{A}}^+ = \lambda_1(\hat{A})$. Now Theorem 1.18 implies that $\mathcal{S}(A) = \mathcal{S}(B) \cup \mathcal{S}(D)$ and thus $\lambda_1(A) = \max(\lambda_1(B), \lambda_1(D))$. Hence the result. □

Example 5.1. *We describe completely the two dimensional case. Let $(A^i)_{i\in E}$ be a family of 2×2 upper triangular matrices :*

$$A^i = \begin{pmatrix} b_i & 0 \\ c_i & d_i \end{pmatrix}.$$

One has $\Lambda_B^+ = \Lambda_B^- = \sum_i p_i b_i := \Lambda_B$ and $\Lambda_D^+ = \Lambda_D^- = \sum_i p_i d_i := \Lambda_D$. We have the following :

1. *If $\Lambda_B > \Lambda_D$, then $\Lambda_A^+ = \Lambda_B$ and $\Lambda_A^- = \Lambda_D$;*
2. *If $\Lambda_B = \Lambda_D$, then $\Lambda_A^+ = \Lambda_A^- = \Lambda_B = \Lambda_D$;*
3. *If for all $i \neq j$, $c_i(b_j - d_j) = c_j(b_i - d_i)$, then $\Lambda_A^+ = \max(\Lambda_B, \Lambda_D)$ and $\Lambda_A^- = \min(\Lambda_B, \Lambda_D)$;*
4. *If $\Lambda_B < \Lambda_D$ and if there exist $i \neq j$ such that $c_i(b_j - d_j) \neq c_j(b_i - d_i)$, then $\Lambda_A^+ = \Lambda_A^- = \Lambda_D$.*

In case where $\Lambda_B > \Lambda_D$, then $\Lambda_A^+ = \Lambda_B$ by Proposition 5.2 and since by Lemma 5.1, $\Lambda(D) \subset \Lambda(A)$, one has $\Lambda_A^- = \Lambda_D$.

If $\Lambda_B = \Lambda_D$, the set of Lyapunov exponents of $(A^i)_{i\in E}$ in the sense of ergodic theory reduces to Λ_B, hence the result (see proof of Proposition 5.1 in Section 5.4).

Now assume that for all $i \neq j$, $c_i(b_j - d_j) = c_j(b_i - d_i)$. If $\Lambda_B = \Lambda_D$, then the result follows from point 2. If $\Lambda_B \neq \Lambda_D$, there exists $i_0 \in E$ such that $b_{i_0} \neq d_{i_0}$. Set $x^* = (1, \frac{c_{i_0}}{b_{i_0} - d_{i_0}})$ We claim that x^* is a common eigenvector for all the A^i. Indeed, let $i \in E$. If $b_i - d_i \neq 0$, then $x = (1, \frac{c_i}{b_i - d_i})$ is an eigenvector of A^i associated with b_i, and since $c_i(b_{i_0} - d_{i_0}) = c_{i_0}(b_i - d_i)$, $x = x^*$. If $b_i - d_i = 0$, since $b_{i_0} - d_{i_0} \neq 0$ and $c_i(b_{i_0} - d_{i_0}) = c_{i_0}(b_i - d_i)$, one has $c_i = 0$, in other words $A^i = b_i I$, hence x^* is an eigenvector of A^i. We conclude that if we let $\theta^* = \frac{x^*}{\|x^*\|}$, then $\mu = \delta_{\theta^*} \otimes p \in \mathcal{P}_{inv}^{(\Theta(A),J)}$ and $\Lambda_A(\mu) = \Lambda_B$. This combined with Lemma 5.1 proves point 3.

Finally assume that $\Lambda_B < \Lambda_D$. It implies that there exists $i_0 \in E$ such that $b_{i_0} < d_{i_0}$. Thus, $\{(0,1), (0-1)\}$ is accessible from every point in $S^{d-1} \setminus \{\theta^*\}$ where θ^* is defined as before. Now there exist $j \in E$ such that θ^* is not an eigenvector for A^j. In particular, we can reach $S^{d-1} \setminus \{\theta^*\}$ from θ^* by following A^j. Hence $\{(0,1), (0-1)\}$ is accessible from S^{d-1} and the result follows from Proposition 5.1.

5.2.3 Main Results

The first theorem is an immediate consequence of Proposition 5.2 and Theorem 3.1.

Theorem 5.1. *Assume $\Lambda_B^+ < 0$ and $\Lambda_D^+ < 0$. Let $0 < \alpha < -\Lambda_A^+$. Then there exists a neighborhood $\mathcal{U}$ of 0 and $\eta > 0$ such that for all $x \in \mathcal{U}$ and $i \in E$*

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

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

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

The next theorem is the main result of this paper. It gives results when the Lyapunov exponents are of opposite signs. 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 $A \subset M \times E$

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

is then the proportion of the time spent by Z in A up to time t. Recall that $M_+ = \{(x_n, x_m) \in M \ : \ x_n \neq 0\}$, and for $\delta > 0$, set $M_0^\delta = \{(x_n, x_m) \in M_+ \ : \ \|x_n\| < \delta\}$.

Theorem 5.2. *Assume $\Lambda_B^- > 0 > \Lambda_D^+$ and that 0 is accessible from $M_0 \times E$. Then :*

(i) *For all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x \in M_+$, $i \in E$, $\mathbb{P}_{x,i}$ almost surely,*

$$\limsup_{t\to\infty} \Pi_t(M_0^\delta \times E) \leq \varepsilon.$$

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

(ii) *There exist positive constants θ, K such that for all $\mu \in \mathcal{P}_{inv}(M_+ \times E)$*

$$\sum_{i\in E} \int \|x_n\|^{-\theta} d\mu^i(x_n, x_m) \leq K.$$

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

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

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

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

Remark 5.2. Note that under the assumptions of the above theorem, Lemma 5.1 implies $\Lambda_A^- \leq \Lambda_D^- < 0$ while by Proposition 5.2, $\Lambda_A^+ = \Lambda_B^+ > 0$. Thus the results in Chapter 3 cannot be applied.

As in Chapter 3, we give the following theorem ensuring uniqueness of the invariant probability giving no mass to $M_0 \times E$.

Theorem 5.3. *In addition to the assumptions of Theorem 5.2, suppose that there exists a point $y \in M_+$ accessible from M_+ at which the weak bracket condition holds. Then*

(i) *The set $\mathcal{P}^Z_{inv}(M_+ \times E)$ reduces to a single element, denoted Π;*

(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.

Strengthening the bracket condition leads to the following result.

Theorem 5.4. *In addition to the assumptions of Theorem 5.2, suppose that one of the two following holds :*

(i) *The weak bracket condition is strengthened to the strong bracket condition; or*

(ii) *There exist $\alpha_1, \ldots, \alpha_N \in \mathbb{R}$ with $\sum \alpha_i = 1$ and a point $e^\star \in M_+$ accessible from M_+ such that $\sum \alpha_i F^i(e^\star) = 0$.*

Then there exist $\kappa, \theta > 0$ such that for all $x \in M_+$ and $i \in E$,

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

5.3 Examples

In this section we give several examples of applications of our results.

5.3.1 Lorenz Vector Fields

In [BH12], the authors consider a random switching between two Lorenz vector fields F^i, $i = 0, 1$:

$$F^i(x, y, z) = \begin{pmatrix} \sigma_i(y - x) \\ r_i x - y - xz \\ xy - b_i z \end{pmatrix}, \tag{5.7}$$

with $\sigma_0 = \sigma_1 = 10$, $b_0 = b_1 = 8/3$, $r_0 = 28$, and $r_1 \neq r_0$ close to 28. It is known since the proof of Tucker [Tuc99] that F^0 admits a robust strange attractor Γ_0. Thus for r_1 close to r_0, F^1 shares this property. In [BH12], it has been shown that the compact set $M = \{(x, y, z) \in \mathbb{R}^3 \ : \ 2r_0\sigma(x^2 + y^2) + 2\sigma b(z_0 - r_0)^2 \leq 2\sigma b r_0^2\}$ is forward invariant, and that Γ_0 is accessible from every point that does not lie on the z-axis. Moreover they proved that the strong bracket condition holds at any point which is not on the z-axis. Then they argue that by compactness of M, there exists an invariant probability, and that it has to be absolutely continuous with respect to the Lebesgue measure due to the bracket condition. However, this argument is not sufficient : there exists indeed an invariant probability measure on M, which is $\delta_0 \otimes p$. However, this measure is not absolutely continuous. We explain how our results apply to that situation and fill in this gap in the proof of [BH12]. In particular, we prove the following result :

Proposition 5.3. *Let F^i, $i = 0, 1$ be two Lorenz vector fields defined by* (5.7) *with $\sigma_0 = \sigma_1 = 10$, $b_0 = b_1 = 8/3$, $r_0 = 28$, and $r_1 \neq r_0$ close to* 28*. Then the process Z admits a unique invariant probability measure Π such that $\Pi(M \setminus \{x = y = 0\}) = 1$. Moreover, Π is absolutely continuous with respect to the Lebesgue measure, and there exist $\kappa, \theta > 0$ such that for all $x_0 = (x, y, z) \in M$ such that $(x, y) \neq 0$ and $i \in E$,*

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

Proof One can note that the z-axis is invariant, and that assumption 5.1 holds with $n = 2$ and $m = 1$. Moreover, we have

$$A^i = \begin{pmatrix} B^i & 0 \\ 0 & -b_i \end{pmatrix}, \tag{5.8}$$

where

$$B^i = \begin{pmatrix} -\sigma_i & \sigma_i \\ r_i & -1 \end{pmatrix}.$$

Setting $D^i = (-b_i)$, one has $\Lambda^+_D = -(p_0 b_0 + p_1 b_1) < 0$. Furthermore, on the z - axis, $|z_t| \leq z_0 e^{-bt}$, with $b = \min(b_0, b_1)$. Hence 0 is accessible from M_0. Let us show that $\Lambda^-_B > 0$. First, it is easily checked that B^0 and B^1 have no common eigenvectors. Therefore, Example 3.2 implies that $\Lambda^+_B = \Lambda^-_B$. Next, the matrices B^i are Metzler, meaning that their off diagonal entries are nonnegative. Therefore, the Kolotilina-type lower estimate for the top Lyapunov exponent proved by Mierczyński [Mie15, Theorem 1.3] implies that

$$\Lambda^-_B \geq \frac{1}{2} \sum_i p_i Tr(B^i) + \sum_i p_i \sqrt{B^i_{12} B^i_{21}}.$$

Here, $Tr(B^i) = -11$ for $i = 0, 1$, and $\sqrt{B^0_{12} B^0_{21}} = \sqrt{280} > 11/2$. Since r_1 is close to r_0, we also have that $\sqrt{B^1_{12} B^1_{21}} > 11/2$, hence $\Lambda^-_B > 0$. The result follows from Theorem 5.4 due to the strong bracket condition proved in [BH12]. □

In Figure 5.1, we show a trajectory of X_t with initial condition $(0, 0.05, 0.05)$ for the vector fields F^0 and F^1 given by the above values of parameters and $r_1 = 35$. The z-axis is drawn in black.

5.3.2 Epidemiological SIRS models

In this section, we show how our result enables to recover and extend those found in [LLC17]. In this paper, the following SIRS model with random switching is studied :

$$F^k(S, I, R) = \begin{pmatrix} \Lambda - \mu S + \lambda_k R - \beta_k S G_k(I) \\ \beta_k S G_k(I) - (\mu + \alpha_k + \delta_k) I \\ \delta_k I - (\mu + \lambda_k) R \end{pmatrix}, \tag{5.9}$$

for $k \in E = \{1, \dots, N\}$, where G_k is a regular function such that $G_k(0) = 0$. The reader is referred to [LLC17] for the epidemiological interpretation of the different constants. The authors study the specific case where only β is allowed to depend on k and where the discrete component $(r_t)_{t \geq 0}$ is an irreducible Markov chain on E, that is the rate matrix a does not depend on the position. Here we assume that the positive constants $\lambda_k, \alpha_k, \delta_k$

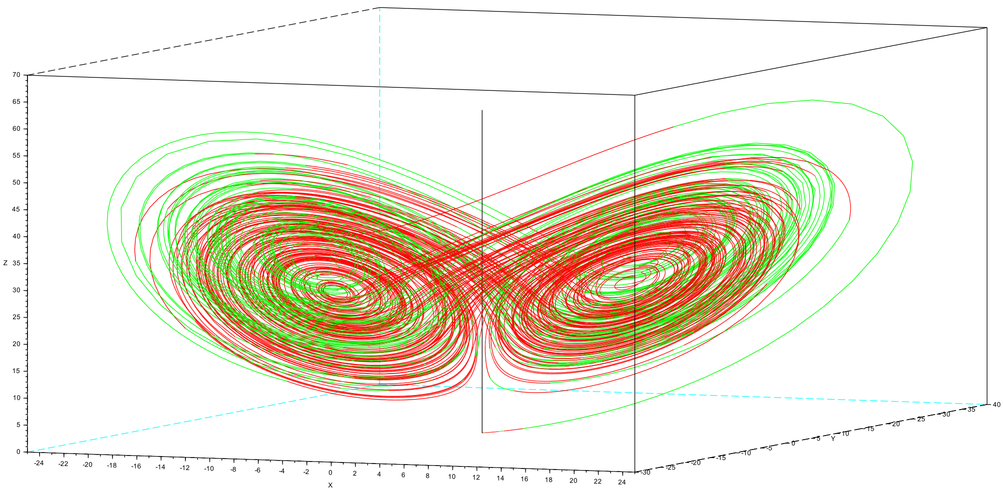


Figure 5.1: Randomly switched Lorenz vector fields

and the functions G_k may depends on k and that a could depend on the position. We still let Λ and μ be constant : they are the intrinsic birth and death rates, and are not related on how the disease spread among the population. Thus the point $q = (\frac{\Lambda}{\mu}, 0, 0)$ is a common equilibrium for the F^k. Set $(Z_t)_{t\geq 0} = (X_t, r_t)_{t\geq 0}$, with $X_t = (S_t, I_t, R_t)$. Write $\mathbb{R}^3_+ = \{x \in \mathbb{R}^3 \ : x_i \geq 0, i = 1, 2, 3\}$ and $\mathbb{R}^3_{++} = \{x \in \mathbb{R}^3 \ : x_i > 0, i = 1, 2, 3\}$. It is easily seen that $\mathbb{R}^3_+$ and $\mathbb{R}^3_{++}$ are positively invariant for all the F^k. Moreover, one can check that the compact set

$$M = \{x \in \mathbb{R}^3_+ \ : \ x_1 + x_2 + x_3 \leq \Lambda/\mu\}$$

is also positively invariant for all the F^k. Furthermore, there are two invariant sets : the S-axis and the (S, R) - plane. We set $M_0 = \{(S, I, R) \in M \ : \ I = 0\}$. We make the following assumptions, that are taken from [LLC17] :

Assumption 5.2.

(i) *For all k, $G_k : \mathbb{R}_+ \to \mathbb{R}_+$ is C^2, with $G_k(0) = 0$ and $0 < G_k(I) \leq G_k'(0)I$ for $I > 0$;*

(ii) *For all k, if $\beta_k \frac{\Lambda}{\mu} G_k'(0) - (\mu + \alpha_k + \delta_k) > 0$, then F^k admits an equilibrium point $x^* \in M_+$ which is accessible from M_+.*

For convenience, we reorder the coordinates as (I, R, S) and set $q = (0, 0, \frac{\Lambda}{\mu})$. Writing $A^k = DF^k(q)$, one has

$$A^k = \begin{pmatrix} \beta_k \frac{\Lambda}{\mu} G_k'(0) - (\mu + \alpha_k + \delta_k) & 0 & 0 \\ \delta & -(\mu + \lambda_k) & 0 \\ -\beta_k \frac{\Lambda}{\mu} G_k'(0) & \lambda_k & -\mu \end{pmatrix}.$$

If we denote by D^k the matrix

$$D^k = \begin{pmatrix} -(\mu + \lambda_k) & 0 \\ \lambda_k & -\mu \end{pmatrix},$$

then by Proposition 5.2, $\Lambda_D^+ = \max(\Lambda_1, \Lambda_2)$, with

$$\Lambda_1 = -\sum_k p_k(\mu + \lambda_k) < 0$$

and

$$\Lambda_2 = -\sum_k p_k\mu = -\mu < 0.$$

Hence $\Lambda_D^+ = -\mu < 0$, and by Theorem 5.1, on M_0, the process converges to q. Now if $B^k = (\beta_k \frac{\Lambda}{\mu} G'(0) - (\mu + \alpha_k + \delta_k))$, then $\Lambda_B^- = \Lambda_B^+ = \sum_k p_k(\beta_k \frac{\Lambda}{\mu} G_k'(0) - (\mu + \alpha_k + \delta_k))$. As in [LLC17], we set

$$R_0 = \frac{\sum_k p_k \beta_k \frac{\Lambda}{\mu} G'(0)}{\sum_k p_k(\mu + \alpha_k + \delta_k)}.$$

Note that $R_0 < 1$ (respectively $R_0 > 1$) if and only if $\Lambda_B^- < 0$ (resp. $\Lambda_B^- > 0$). In particular, Theorems 5.1, 5.2, 5.3 and 5.4 imply the following statement, that recovers and slightly extends Theorems 4, 8 and 9 in [LLC17].

Theorem 5.5. *With the above notation, the following hold.*

(i) *Assume that $R_0 < 1$. Then, for all $z_0 = (s_0, i_0, r_0, k_0) \in M \times E$, one has*

$$\mathbb{P}_{z_0}(\limsup_{t\to\infty} \frac{1}{t} \log(\|(S_t, I_t, R_t) - (\frac{\Lambda}{\mu}, 0, 0)\|) \leq \Lambda_A^+) = 1,$$

where $\Lambda_A^+ = \max(\Lambda_B^+, -\mu)$.

(ii) *Assume that $R_0 > 1$. Then the process Z admits an invariant probability measure Π such that $\Pi(M \setminus M_0^1 \times E) = 1$.*

(iii) *Assume in addition to $R_0 > 1$ that the weak bracket condition holds at an accessible point. Then Π is unique and there exist $\kappa, \theta > 0$ such that for all $x = (s, i, r) \in M_+$ and $k \in E$,*

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

In addition, for all $x \in M_+$ and $k \in E$,

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

$\mathbb{P}_{x,k}$ almost surely.

Proof If $R_0 < 1$, then $\Lambda_B^+ < 0$ and thus there exists $k_0 \in E$ such that $\beta_{k_0} \frac{\Lambda}{\mu} G_{k_0}'(0) - (\mu + \alpha_{k_0} + \delta_{k_0}) < 0$. We show that this implies that q is accessible from M_+. Let $x_0 \in M_+$ and denote by $x_t = (s_t, i_t, r_t)$ the solution of

$$\frac{dx_t}{dt} = F^{k_0}(x_t)$$

with initial condition x_0. Now by assumption 5.2 and the fact that $s_t \leq \frac{\Lambda}{\mu}$,

$$\frac{di_t}{dt} \leq \left(\beta_{k_0} \frac{\Lambda}{\mu} G_{k_0}'(0) - (\mu + \alpha_{k_0} + \delta_{k_0})\right) i_t.$$

Since $\beta_{k_0}\frac{\Lambda}{\mu}G'_{k_0}(0) - (\mu + \alpha_{k_0} + \delta_{k_0}) < 0$, i_t converges to 0 exponentially fast. It is easy to check that on M_0, (s_t, r_t) converges to $(\frac{\Lambda}{\mu}, 0)$, thus x_t converges to q. Hence q is accessible, and **(i)** follows from Theorem 5.1. Point **(ii)** is an immediate consequence of Theorem 5.2. Now if $R_0 > 1$, there exists $k_0 \in E$ such that $\beta_{k_0}\frac{\Lambda}{\mu}G'_{k_0}(0) - (\mu + \alpha_{k_0} + \delta_{k_0}) > 0$. By assumption 5.2, this implies that F^{k_0} admits an accessible equilibrium $x^* \in M_+$. Point **(iii)** follows then by Theorems 5.3 and 5.4. □

5.4 Proofs

5.4.1 Proof of Theorem 5.2

The idea of the proof is similar to that used in Chapter 3, and also relies on results of [Ben18]. In Chapter 3, we rewrite the process in spheric coordinates on $\mathbb{R}_+ \times S^{d-1}$. Here the idea is to only write the spheric coordinates for the part of X_t living in $\mathbb{R}^n$. That is, we consider the map $\Psi : \mathbb{R}^n \setminus \{0_n\} \times \mathbb{R}^m \times E \to \mathbb{R}_+^* \times S^{n-1} \times \mathbb{R}^m \times E$ defined by $\Psi(x_n, x_m, i) = (\|x_n\|, \frac{x_n}{\|x_n\|}, x_m, i)$. We set $\mathcal{X}_+ = \Psi(M_+ \times E)$. When $(x, i) \in M_+ \times E$, the process $\tilde{Z}_t = \Psi(Z_t) = (\rho_t, \Theta_t, X_t^m, I_t)$ is well defined and satisfies

$$\begin{cases} \frac{d\rho_t}{dt} = \langle \Theta_t, \tilde{F}_n^{I_t}(\rho_t, \Theta_t, X_t^m)\rangle \rho_t \\ \\ \frac{d\Theta_t}{dt} = \tilde{F}_n^{I_t}(\rho_t, \Theta_t, X_t^m) - \langle \Theta_t, \tilde{F}_n^{I_t}(\rho_t, \Theta_t, X_t^m)\rangle \Theta_t \\ \\ \frac{dX_t^m}{dt} = \tilde{F}_m^{I_t}(\rho_t, \Theta_t, X_t^m) \\ \\ \mathsf{P}(I_{t+s} = j | \mathcal{F}_t) = a_{ij}(\rho_t\Theta_t, X_t^m)s + o(s) \text{ for } i \neq j \text{ on } \{I_t = i\} \end{cases} \tag{5.10}$$

where for all $(\rho, \theta, x_m) \in \mathbb{R}_+^* \times S^{n-1} \times \mathbb{R}^m$, $\tilde{F}_m^i(\rho, \theta, x_m) = F_m^i(\rho\theta, x_m)$ and

$$\tilde{F}_n^i(\rho, \theta, x_m) = \frac{F_n^i(\rho\theta, x_m)}{\rho}.$$

Since F_n^i is C^2 and $F_n^i(0, x_m) = 0$, the map $\tilde{F}_n^i$ extends to a C^1 map on $\mathbb{R}_+ \times S^{n-1} \times \mathbb{R}^m$ by setting

$$\tilde{F}_n^i(0, \theta, x_m) = B^i(x_m)\theta,$$

where $B^i(x_m) \in M_n(\mathbb{R})$ is such that $DF_n^i(0, x_m) = (B^i(x_m), 0)$. Note that in particular, $B^i(0_m) = B^i$. Thanks to this definition, we can extend (5.10) to

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

where $\mathcal{X}_0 = \{0\} \times S^{n-1} \times \mathbb{R}^m \times E$. This induces a PDMP (still denoted $\tilde{Z}$) on $\mathcal{X}$, whose infinitesimal generator $\tilde{\mathcal{L}}$ acts on functions $f : \mathcal{X} \to \mathbb{R}$ smooth in (ρ, θ, x_m) according to

$$\begin{aligned} \tilde{\mathcal{L}}f(\rho, \theta, x_m, i) = & \frac{\partial f^i}{\partial \rho}(\rho, \theta, x_m)\langle \theta, \tilde{F}_n^i(\rho, \theta, x_m)\rangle \rho + \langle \nabla_\theta f^i(\rho, \theta, x_m), \tilde{G}^i(\rho, \theta, x_m)\rangle \\ & + \langle \nabla_{x_m} f^i(\rho, \theta, x_m), \tilde{F}_m^i(\rho, \theta, x_m)\rangle \\ & + \sum_{j \in E} a_{ij}(\rho\theta, x_m)(f^j(\rho, \theta, x_m) - f^i(\rho, \theta, x_m)), \end{aligned} \tag{5.11}$$

where $\tilde{G}^i(\rho,\theta,x_m) = \tilde{F}^i_n(\rho,\theta,x_m) - \langle\theta, \tilde{F}^i_n(\rho,\theta,x_m)\rangle\theta$.

The set $\mathcal{X}_0$ is invariant, and we identify it with $S^{n-1}\times\mathbb{R}^m\times E$. On this set, the process (Θ, X^m, I) satisfies

$$\begin{cases} \frac{d\Theta_t}{dt} = B^{I_t}(X^m_t)\Theta_t - \langle\Theta_t, B^{I_t}(X^m_t)\Theta_t\rangle\Theta_t \\ \frac{dX^m_t}{dt} = F^{I_t}_m(0, X^m_t) \\ \mathsf{P}(I_{t+s} = j|\mathcal{F}_t) = a_{ij}(0,X^m_t)s + o(s) \text{ for } i\neq j \text{ on } \{I_t = i\} \end{cases} \tag{5.12}$$

Lemma 5.2. *For all $(\theta, x_m, i)\in\mathcal{X}_0$, one has*

$$\mathbb{P}_{\theta,x_m,i}(\limsup_{t\to\infty}\frac{1}{t}\log(\|X^m_t\|)\leq\Lambda^+_D) = 1.$$

Proof On $\mathcal{X}_0$, the process (X^m, I) evolves independently from Θ. It is a PDMP with vector fields $\hat{F}^i:\mathbb{R}^m\to\mathbb{R}^m$ and transition rate matrix $(\hat{a}_{ij})$ defined for all $x\in\mathbb{R}^m$ respectively by $\hat{F}^i(x) = F^i_m(0_n, x)$ and $\hat{a}_{ij}(x) = a_{ij}(0_n,x)$. The origin 0_m is a common zero for all the $\hat{F}^i$, and $D\hat{F}^i(0_m) = D^i$. In particular, the maximal Lyapunov exponent for (X^m, I) is Λ^+_D and the result follows from Theorem 3.1 due to the fact that $\Lambda^+_D<0$ and 0 is accessible from M_0. □

Note that on $\{0\}\times S^{n-1}\times\{0_m\}\times E$, (Θ, I) is equal to the PDMP $(\Theta(B), J)$ defined in section 5.2.2. Therefore, we have :

Lemma 5.3. *Let μ be an invariant probability of $\tilde{Z}$ on $\mathcal{X}_0$. Then $\mu(\mathrm{d}\theta,\mathrm{d}x,\mathrm{d}i) = \delta_0(\mathrm{d}x)\otimes\hat{\mu}(\mathrm{d}\theta,\mathrm{d}i)$ where $\hat{\mu}$ is an invariant probability of $(\Theta(B), J)$.*

Proof Let $(Q_t)_{t\geq0}$ be the semigroup of (Θ, X^m, I) on $\mathcal{X}_0$. Let $f:\mathbb{R}^m\to\mathbb{R}$ be a continuous bounded function and define $\hat{f}:\mathcal{X}_0\to\mathbb{R}$ by $\hat{f}(\theta,x,i) = f(x)$. By invariance of μ, $\mu Q_t\hat{f} = \mu\hat{f}$ for all $t\geq0$. Now, $\mu\hat{f} = \tilde{\mu}f$ where $\tilde{\mu}$ is the marginal of μ on $\mathbb{R}^m$ and by Lemma 5.2 and dominated converge, $\mu Q_t\hat{f}\to f(0)$ when $t\to\infty$. Thus $\tilde{\mu} = \delta_0$. Since the marginal law is a Dirac mass, this implies that μ is a product measure : $\mu = \delta_0\otimes\hat{\mu}$, where $\hat{\mu}$ is the marginal of μ on $S^{n-1}\times E$. The result follows from the remark preceding this lemma. □

Define $H:\mathcal{X}\to\mathbb{R}$ by $H(\rho,\theta,x_m,i) = -\langle\theta,\tilde{F}^i_n(\rho,\theta,x_m)\rangle$. The following lemma is immediate from Lemma 5.3 and the definition of H.

Lemma 5.4. *Let μ be an invariant probability on $\mathcal{X}_0$. Then with the notation of Lemma (5.3), $\mu H = -\Lambda_B(\hat{\mu})$.*

Now we proceed to the proof of Theorem 5.2. Letting $V:\mathcal{X}_+\to\mathbb{R}_+$ be a smooth function coinciding with $-\log(\rho)$ for all $(\rho,\theta,x_m,i)\in\mathcal{X}$ such that $\rho\leq1$, the end of the proof is verbatim the same as in Section 3.5 of Chapter 3 by noting that $\tilde{\mathcal{L}}V = H$. □

5.4.2 Proof of Proposition 5.1

The proof is really similar to the one of Theorem 5.2, so we do not give all the details. Recall from proof of Lemma 5.1 that we rewrite (5.3) as

$$\begin{cases} \frac{d\Theta_t^n}{dt} = B^{J_t}\Theta_t^n - \left(\langle B^{J_t}\Theta_t^n, \Theta_t^n\rangle + \langle C^{J_t}\Theta_t^n + D^{J_t}\Theta_t^m, \Theta_t^m\rangle\right)\Theta_t^n \\ \frac{d\Theta_t^m}{dt} = C^{J_t}\Theta_t^n + D^{J_t}\Theta_t^m - \left(\langle B^{J_t}\Theta_t^n, \Theta_t^n\rangle + \langle C^{J_t}\Theta_t^n + D^{J_t}\Theta_t^m, \Theta_t^m\rangle\right)\Theta_t^m \end{cases}$$

As in the proof of Theorem 5.2, we write $\Theta^n = \rho\hat{\Theta}$ with $\rho = \|\Theta^n\|$ and $\hat{\Theta} = \frac{\Theta^n}{\rho} \in S^{n-1}$. As before, the set $\{\rho = 0\}$ is invariant, and one can check that on this state, $\hat{\Theta}$ and Θ^m evolves independently as :

$$\begin{cases} \frac{d\hat{\Theta}_t}{dt} = B^{J_t}\hat{\Theta} - \langle B^{J_t}\hat{\Theta}, \hat{\Theta}\rangle\hat{\Theta} \\ \frac{d\Theta_t^m}{dt} = D^{J_t}\Theta_t^m - \langle D^{J_t}\Theta_t^m, \Theta_t^m\rangle\Theta_t^m \end{cases} \tag{5.13}$$

That is $(\hat{\Theta}, J) = (\Theta(B), J)$ and $(\Theta^m, J) = (\Theta(D), J)$. Furthermore, setting $\hat{V}(\rho, \hat{\theta}, \theta^m, i) = -\log(\rho)$, one has

$$\hat{L}\hat{V}(\rho, \hat{\theta}, \theta^m, i) = -\langle B^i\hat{\theta}, \hat{\theta}\rangle + \left(\rho^2\langle B^i\hat{\theta}, \hat{\theta}\rangle + \rho\langle C^i\hat{\theta}, \theta^m\rangle + \langle D^i\theta^m, \theta^m\rangle\right) := \hat{H}(\rho, \hat{\theta}, \theta^m, i).$$

Here $\hat{L}$ stands for the generator of $\hat{Z} := (\rho, \hat{\Theta}, \Theta^m, J)$. Now if μ is an invariant probability of $\hat{Z}$ on $\{\rho = 0\}$; then there are $\hat{\mu} \in \mathcal{P}_{inv}^{(\Theta(B),J)}$ and $\tilde{\mu} \in \mathcal{P}_{inv}^{(\Theta(D),J)}$ such that $\mu\hat{H} = -\Lambda_B(\hat{\mu}) + \Lambda_D(\tilde{\mu})$. In particular, if $\Lambda_B^+ < \Lambda_D^-$, then for all $\mu \in \mathcal{P}_{inv}^{\hat{Z}}$ with $\mu(\{\rho = 0\}) = 1$, one has $\mu\hat{H} > 0$. Moreover, since we assumed that $\{(\theta_n, \theta_m) \in S^{d-1} : \theta_n = 0\}$ is accessible from S for $(\Theta(A), J)$, the set $\{\rho = 0\}$ is accessible for $\hat{Z}$. This concludes the proof by the same arguments as in Chapter 3.

Remark 5.3. The same proof shows that if $\Lambda_B^- > \Lambda_D^+$, $(\Theta(A), J)$ admits at least one invariant probability measure giving no mass to $\{(\theta_n, \theta_m) \in S^{d-1} : \theta_n = 0\}$. An interesting question would be to know if it is possible to recover the Lyapunov exponents associated to B with this invariant probability measure, like in dimension 2 (see Example 5.1.)

Chapter 6

Approximation with extinction of Markov processes that never die

Let $(X_t)_{t\geq 0}$ be a Markov process on a compact set, leaving invariant a closed subset M_0, seen as the extinction set. We consider a family of Markov processes $(X^N)_{N\geq 0}$ that converges to X in probability on any finite time interval, with the main difference that while X cannot dies in finite time, X^N gets extinct almost surely in finite time. In this chapter, we study the behaviour of the family of quasi-stationary distributions α^N associated with X^N. Using the stochastic persistence theory developed in [Ben18], we show that when X is persistent, every limit point of α^N is an invariant probability for X, and, under some assumption, that this probability does not give mass to M_0. In that case, the mean extinction time of X^N goes to infinity when N goes. On the other hand, when X dies on a infinite horizon, every limit point of α^N is concentrated on the extinction set. We give applications to an epidemic model in random environment that converges to a Piecewise Deterministic Markov Process (PDMP).

Keywords: Persistence, Quasi-stationary distribution, Piecewise deterministic Markov processes; Epidemic models; SIS

MSC primary: 60K35, 60G17, 60J60

Contents

6.1 Introduction

Numerous models in ecology are represented by the solution of an Ordinary Differential Equation (ODE). While these models enable us, more or less easily, to understand some behaviours observed in Nature, they do not take into account two specific matters inherent to real life : randomness in the environment and the tragic destiny of every population - death in finite time. Nonetheless, it is now well known that such ODEs may appear as limits of finite population models described by finite Markov Chains, when the size of the population goes to infinity (see the work of Kurtz, especially [Kur81]). This gives a clue on why we are very unlikely to observe the extinction of a large population in the real world, yet we know that it will inevitably occurs. Instead of that, we have better chance to see the population reaching a metastable equilibrium, which can be related, if they exist, to the quasi-stationary distributions (QSD) of the finite size Markov Chain. A natural question is then to ask what is the behaviour of this family of QSDs when N, the size of the population, goes to infinity. In [FS14], Faure and Schreiber study this problem in the discrete time setting, and show that when the limit process has an attractor bounded away from the extinction set, then the QSDs concentrate on this attractor when N gets large. Moreover, they proved that the mean extinction time starting from the QSDs goes to infinity exponentially fast with the size of the population. On the contrary, if the extinction set is an attractor of the limit system, then the QSDs will eventually concentrate on that set. These results have been slightly generalized in unpublished works by Marmet [Mar13], as well as in [Mar14] for diffusion type approximations. More recently, Chazottes, Collet and Méléard study the limit of a Birth-and-death process and get some sharp asymptotics for the extinction rate [CCM16] and [CCM17]. In the particular case of dimension 1, they show that, as the parameter goes to infinity, the QSD of the birth and death process approaches a Gaussian law centred on the unique equilibrium of the limiting system.

Our goal in this chapter is to give some results when the limit system is no longer an ODE but a Markov process that never dies in finite time. In particular, contrary to the aforementioned studies, the randomness is still present in the asymptotic process. Thus, the dynamic might be more intricate, and we will need some tools to catch the behaviour of that Markov process. More precisely, we will use the theory of stochastic persistence developed by Benaïm in [Ben18], and inspired by pioneer works of Hofbauer and Schreiber. Briefly put, our results states that :

- If the limiting process X is persistent, then every limit point of the QSD is an invariant probability of X, and the extinction rate goes to 0.
- If the limiting process X is non persistent, then every limit point of the QSD put all its mass on the extinction set.

The chapter is organised as follows. In Section 6.2, the framework and assumptions are given. The main results are stated precisely in Section 6.3. In section 6.4, we give an application to an epidemic birth-and-death process in random environment, whose scaling limit is a Piecewise Deterministic Markov Process. Section 6.5 is devoted to the proofs.

6.2 Assumptions

In this section we present our assumptions. We consider a continuous-time Markov process $(X_t)_{t\geq 0}$ defined on some measurable space $(\Omega, \mathcal{F})$ with values in a metric space (M, d).

Assumption 6.1. *(M,d) is a compact metric space.*

We consider, for all $N \in \mathbb{N}$, a càdlàg Markov process $(X_t^N)_{t\geq 0}$ defined on the same measurable space $(\Omega, \mathcal{F})$, living on some subset M^N of M and enjoying the following properties:

Assumption 6.2 (Absorbing set)**.** *There exists a closed set $M_0 \subset M$ such that, for all N:*

(i) *M_0 is absorbing : for all $t,s \geq 0$, $X_t^N \in M_0 \Rightarrow X_{t+s}^N \in M_0$;*

(ii) *For all $x \in M^N$, $\mathbb{P}_x(T_0^N < \infty) = 1$, where $T_0^N = \inf\{t \geq 0 \,:\, X_t^N \in M_0\}$;*

(iii) *For all $x \in M^N$ and $t \geq 0$, $\mathbb{P}_x(T_0^N > t) > 0$.*

The set $M_0^N = M_0 \cap M^N$ will be seen as the *extinction set* of the processes X^N. We set $M_+ = M \setminus M_0$ and $M_+^N = M_+ \cap M^N$. Recall that for a Markov process Z satisfying Assumption 6.2, a *quasi-stationary distribution* (QSD) is a probability distribution α on M_+ such that, for all $A \in \mathcal{B}(M_+)$ and all $t \geq 0$,

$$\mathbb{P}_\alpha(Z_t \in A | T_0 > t) = \alpha(A).$$

For $\delta > 0$, we set $M_{+,\delta} = \{x \in M \,:\, d(x, M_0) > \delta\}$.

Assumption 6.3 (Quasi-stationary distribution)**.** *For all N, the process X^N admits a quasi-stationary distribution α^N. Moreover, there exist $\delta > 0$ such that for N large enough, $\alpha^N(M_{+,\delta}) > 0$.*

The following assumption states that the processes X^N converge to X. :

Assumption 6.4 (Convergence)**.** *For all $T, \delta > 0$,*

$$\lim_{N\to\infty} \max_{x\in M; x^N \to x} \mathbb{P}_{(x^N,x)} \left(\sup_{t\in[0,T]} d(X_t^N, X_t) > \delta \right) = 0. \tag{6.1}$$

The limit in (6.1) as to be understood as follows : for all $\varepsilon > 0$, there exists $N_0 \geq 1$ and $\beta > 0$, such that for all $N \geq N_0$, for all $x \in M$, for all $x^N \in M^N$ such that $d(x^N, x) \leq \beta$, then $\mathbb{P}_{(x^N,x)} \left(\sup_{t\in[0,T]} d(X_t^N, X_t) > \delta\right) \leq \varepsilon$. Here, as always, $\mathbb{P}_{(x^N,x)}$ designs a probability measure on $(\Omega, \mathcal{F})$ such that

$$\mathbb{P}_{(x^N,x)} \left(X_0 = x; X_0^N = x^N\right) = 1.$$

This kind of convergence appears frequently when studying stochastic approximation of ODE, and typically in the works of Kurtz [Kur81] cited in the introduction.

Remark 6.1. *Let F be some smooth vector field on M and let $(\varphi_t)_{t\geq 0}$ denoted the flow induced by the differential equation*

$$\frac{dx_t}{dt} = F(x_t).$$

Then $(\varphi_t)_{t\geq 0}$ can be seen as a Markov process, with semigroup $P_t f(x) = f \circ \varphi_t(x)$. In particular, all the statements of this paper hold when $(X_t)_{t\geq 0}$ is the solution of an ODE.

We now give some information about the limit process :

Assumption 6.5 (Limit process)**.** *The limit process X satisfies the following :*

(i) *The set M_0 (and so M_+) is invariant for X : for all $t \geq 0$, $P_t^X \mathbb{1}_{M_0} = \mathbb{1}_{M_0}$;*

(ii) *The semigroup P^X if C_b -Feller.*

Point **(i)** of Assumption 6.5 has the following consequence : contrary to X^N, the process X can *never die in finite time.*

To shorten notations, we will set P_t^N for $P_t^{X^N}$ and P_t for P_t^X. We will also introduce the semigroup before extinction of X_t^N : for all N, $t \geq 0$ and $x \in M$,

$$\tilde{P}_t^N = \mathbb{E}_x \left(f(X_t^N) \mathbb{1}_{T_0^N > t} \right).$$

We recall the following results on QSD (see Proposition 1.16):

Proposition 6.1. *For all N, there exists a positive number λ_N such that, under α^N, T_0^N has exponential law with parameter λ_N :*

$$\mathbb{P}_{\alpha^N}(T_0^N > t) = e^{-\lambda_N t}.$$

In particular, for all $t \geq 0$,

$$\alpha^N \tilde{P}_t^N f = e^{-\lambda_N t} \alpha^N f. \tag{6.2}$$

Example 6.1 (A toy example)**.** *Let $(X_t)_{t\geq 0}$ be a càdlàg Feller Markov process on some compact subset M of $\mathbb{R}^d$. For all $N \geq 0$, let T^N be an exponential random variable with parameter $\lambda_N > 0$, independent of $(X_t)_{t\geq 0}$. T^N represents the arrival time of a catastrophic event that destroy all the population. Fix some $\partial \in M_0$ and define, for $N \geq 0$, the Markov process X^N by :*

$$X_t^N = \begin{cases} X_t & \text{if} \quad t < T^N \\ \partial & \text{otherwise,} \end{cases}$$

if $x_0 \in M_+$, and $X_t^N = X_t$ if $x_0 \in M_0$. This is a classical construction with the interesting property that $\mathbb{P}_\mu(X_t^N \in A | T^N > t) = \mathbb{P}_\mu(X_t \in A)$ for all $A \in \mathcal{B}(M_+)$ and $\mu \in \mathcal{P}(M_+)$. In particular, α^N is a QSD for X^N if and only if it is an invariant probability measure for X. Moreover, it is clear from the definition of X^N that for all $T > 0$ and $x \in M_+$,

$$\lim_{\delta \to 0} \mathbb{P}_x \left(\max_{t \in [0,T]} d(X_t^N, X_t) > \delta \right) = \mathbb{P}_x(T \geq T^N) = 1 - e^{-\lambda_N T}.$$

This shows that Hypothesis 6.5 holds if and only if λ_N converges to 0. This simple example illustrates the importance of Hypothesis 6.3. In this situation, it holds if and only if X admits an invariant probability measure α such that $\alpha(M_+) = 1$. In that case, $\alpha^N = \alpha$ for all N.

The aim of this chapter is to study the limit points of $(\alpha^N)_{N\geq 0}$, as well as the limit of $(\lambda_N)_{N\geq 0}$. It appears that these limits are related to the behaviour of the limit process X near the extinction set. We will rely on the recent general theory of *stochastic persistence* developed in [Ben18] and exposed in Section 1.3 of Chapter 1. In Section 6.3 we state our main results, that will be proved in Section 6.5. In Section 6.4, we give some examples of Markov Chain in random environment modelling the spread of a disease and converging to a PDMP as the size of the population goes to infinity. We also consider a family of killed Markov processes with a killing rate converging to 0.

6.3 Results

6.3.1 Preliminary results

Throughout this section, α denotes a weak limit point of $(\alpha^N)_{N\geq 0}$. To shorten notation, we assume that α^N converges weakly to α.

Lemma 6.1. *For all $t \geq 0$ and all continuous function f, $\|P_t f - P_t^N f\|_\infty \to 0$.*

Proof By uniform continuity of f, for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x \in M$,

$$|P_t f(x) - P_t^N f(x)| \leq \varepsilon + 2\|f\|_\infty \mathbb{P}_x \left(d(X_t, X_t^N) > \delta\right),$$

hence the result by (6.1). □

Lemma 6.2. *For all $t \geq 0$, for all N, for all $f : M \to \mathbb{R}$ bounded measurable,*

$$|\alpha^N P_t f - e^{-\lambda_N t}\alpha^N f| \leq \|P_t f - P_t^N f\|_\infty + \|f\|_{M_0}(1 - e^{-\lambda_N t}).,\tag{6.3}$$

where $\|f\|_{M_0} = \sup_{x\in M_0} |f(x)|$.

Proof By definition of P_t^N and $\tilde{P}_t^N$, one has for all $x \in M$,

$$|P_t^N f(x) - \tilde{P}_t^N f(x)| \leq \|f\|_{M_0}\mathbb{P}_x(T_0^N \leq t).$$

Integrating this inequality with respect to α^N and by Proposition 6.1, one gets

$$|\alpha^N P_t^N f - \alpha^N \tilde{P}_t^N f| \leq \|f\|_{M_0}(1 - e^{-\lambda_N t}).$$

Now by (6.2), one has

$$|\alpha^N P_t f - e^{-\lambda_N t}\alpha^N f| = |\alpha^N P_t f - \alpha^N \tilde{P}_t^N f|,$$

hence the result by triangular inequality. □

Let $\lambda \in [0, +\infty]$ be a limit point of the sequence $(\lambda_N)_{N\geq 0}$.

Lemma 6.3. *If $\lambda = 0$, then $\alpha \in \mathcal{P}_{inv}$.*

Proof Since $\lambda_N \to 0$, the left-hand side of (6.3) converges to $|\alpha P_t f - \alpha f|$ for all continuous function f. By lemma 6.1, the right-hand side converges to 0. Thus $\alpha P_t f = \alpha f$ for all $t \geq 0$ and all continuous function f, meaning that $\alpha \in \mathcal{P}_{inv}$. □

Lemma 6.4. *For all $t \geq 0$, for all continuous f null on M_0, one has*

$$\alpha P_t f = e^{-\lambda t}\alpha f.\tag{6.4}$$

In particular if $\alpha(M_+) > 0$, then $\lambda = 0$ and hence $\alpha \in \mathcal{P}_{inv}$.

Proof Since $f = 0$ on M_0, (6.4) is a consequence of (6.3). Now, one can find a sequence of continuous nonnegative functions f_n, null on M_0, converging monotonically to $\mathbb{1}_{M_+}$. Thus by (6.4) and monotone convergence, one gets

$$\alpha P_t \mathbb{1}_{M_+} = e^{-\lambda t}\alpha \mathbb{1}_{M_+}.$$

Since M_+ is invariant for P_t, one has $\alpha(M_+) = e^{-\lambda t}\alpha(M_+)$, hence the result. □

6.3.2 Main results

The two following theorems are the main results of this chapter. The first one deals with the persistent case. When we say that X is H-persistent (respectively, H-nonpersistent), we mean that there exist functions (V, H) satisfying Assumption 1.5 with positive (respectively, negative) H-exponents (see Section 1.3 for details).

Theorem 6.1. *Assume X is H-persistent. Then*

$$\lim_{N\to\infty} \lambda^N = 0.$$

In particular, every weak limit point of $(\alpha^N)_{N\geq 0}$ is an invariant probability measure of X.

The second theorem applied when the process X is nonpersistent.

Theorem 6.2. *Assume that for all $\mu \in \mathcal{P}_{inv}$, $\mu(M_0) = 1$. Then $\alpha(M_0) = 1$. In particular, if X is H- nonpersistent and M_0 is accessible, $\alpha(M_0) = 1$.*

Proof If $\lambda = 0$, then $\alpha \in \mathcal{P}_{inv}$ and thus $\alpha(M_0) = 1$ by assumption. If $\lambda > 0$, then $\alpha(M_0) = 1$ by Lemma 6.3. Now if X is H-nonpersistent and M_0 is accessible, by Theorem 1.14, X_t converges almost surely to M_0. In particular, every invariant measure should have its support in M_0. □

6.4 Applications

6.4.1 Approximation of Epidemiological models in random environment

In Chapter 3, we study a SIS model in a randomly fluctuating environment. The resulting process is a Piecewise Deterministic Markov Process (PDMP). In this section we show that we can see this PDMP as a limit of a finite population model.

We consider a finite population of size N, divided into d groups. We assume the number of individuals in each group i is constant and equal to N_i. We assume that, for all i, there exists $p_i \in (0, 1)$ such that

$$\lim_{N\to\infty} \frac{N_i}{N} = p_i.$$

The disease evolves in the population according to the state $k \in E$ of the environment, where E is a finite set. We can describe the evolution of the disease as follows. Each individuals has a random exponential clock with parameter 1, independent from the other ones. The time of the first ring is exponentially distributed with parameter N, and the first ringing clock is carried by an individual of group i with probability N_i/N. If the chosen individual is infected, then she is cured with probability $B_i^k(x) > 0$, where $x = (x_1, \ldots, x_d)$ with $x_i \in \{0, 1/N_i \ldots, 1\}$ the proportion of infected individuals in group i, and k the state of the environment. If she is susceptible, then she becomes infected with probability $C_i^k(x) > 0$. We set $B_i^k(0) = C_i^k(0) = 0$. In parallel, the environment switches to state k to state l with probability $a_{k,l}(x)$. Denote by X_t^N the vector of proportion of infected at time t, and by I_t^N the environment state at time t. Then $(X_t^N, I_t^N)_{t\geq 0}$ is a continuous time Markov Chain on the space

$$M_N = \{0, \frac{1}{N_1}, \dots, 1\} \times \dots \times \{0, \frac{1}{N_d}, \dots, 1\} \times E,$$

with transition rates given for all $(x, k) \in M_N$ by

$$\begin{aligned}(x,k) \longrightarrow &(x + \frac{e_i}{N_i}, k) \quad \text{at rate} \quad N\frac{N_i}{N}B_i^k(x)\\ &(x - \frac{e_i}{N_i}, k) \quad \text{at rate} \quad N\frac{N_i}{N}C_i^k(x)\\ &(x, l) \qquad\qquad \text{at rate} \quad a_{k,l}(x).\end{aligned}$$

Here and throughout, $(e_1, \dots, e_d)$ stands for the canonical basis of $\mathbb{R}^d$. The infinitesimal generator of $Z^N = (X^N, I^N)$ is the operator L^N which acts on bounded measurable functions $f : M_N \to \mathbb{R}$ according to

$$\begin{aligned}L^N f(x,k) = &\sum_{i=1}^d \left[f(x + \frac{e_i}{N_i}, k) - f(x,k) \right] N_i B_i^k(x)\\ &+ \sum_{i=1}^d \left[f(x - \frac{e_i}{N_i}, k) - f(x,k) \right] N_i C_i^k(x)\\ &+ \sum_{l \in E} a_{k,l}(x) \left[f(x,l) - f(x,k) \right].\end{aligned}$$

If f is a smooth enough function, one can see that $L^N f$ converges to the function $L^\infty f$ given by

$$L^\infty f(x,k) = \langle F^k(x,) \nabla f(x,k) \rangle + \sum_{l \in E} a_{k,l}(x) \left[f(x,l) - f(x,k) \right].$$

The operator L^∞ is the generator of the PDMP given above. In [CDMR12], the authors consider reactions in gene network leading to process with generators similar to L^N. They show [CDMR12, Theorem 3.1] that the process Z^N converges in Skorokhod space to the process Z with infinitesimal generator L^∞, provided the initial condition of Z^N converges in distribution to the initial condition of Z. However, this results do not give any information about the quasi-stationary distribution of Z^N. We now show that the process Z^N and Z satisfy the assumptions of section 6.2. Also note that the set $M = [0,1]^d \times E$ with the distance d defined for all $(x,i), (y,j)$ in M by $d((x,i),(y,j)) = \mathbb{1}_{i \neq j} + \mathbb{1}_{i=j}\|x - y\|$ satisfies hypothesis 6.1.

Proposition 6.2. *The process Z^N and Z satisfy hypothesis 6.2 to 6.5.*

In particular, the results of Theorem 6.1 and 6.2 hold.

6.4.2 Process with soft killing

In this section, we study the generalisation of the toy model considered in Example 6.1. That is, we consider a càdlàg Feller Markov process $(X_t)_{t\geq 0}$ on some compact metric space

(M, d), leaving invariant a compact subset M_0. For all $N \geq 0$, let $\kappa_N : M \to \mathbb{R}_+$ be a bounded measurable function, and set

$$T^N = \inf\{t \geq 0 \, : \int_0^t \kappa_N(X_s) ds \geq E\},$$

where E is a random variable, with exponential distribution of parameter one, independent of $(X_t)_{t \geq 0}$. We now define a Markov process $(X_t^N)_{t \geq 0}$ as follows : if $x \in M_0$, then $X_t^N = X_t$ for all $t \geq 0$, if $x \in M_+$,

$$X_t^N = \begin{cases} X_t & \text{if} \quad t < T^N \\ \partial & \text{otherwise}, \end{cases}$$

where ∂ is an arbitrary point in M_0. In order to get convergence (6.1), we impose that $\|\kappa_N\|_\infty$ goes to 0 as N goes to infinity. Since the process X^N lives on the open set M_+ before extinction, there is no reason that without additional assumption, Hypothesis 6.3 is satisfied. Thus, we will assume the following :

Assumption 6.6. *The processes X and X^N satisfy :*

(i) *The process X is H - persistent;*

(ii) *There exists a Doeblin point x^* for X, accessible from M_+;*

(iii) *For all compact set $K \subset M_+$, there exists $c_N(K) > 0$ such that*

$$\sup_{t \geq 0} \frac{\sup_{x \in K} \mathbb{P}_x(t < T^N)}{\inf_{x \in K} \mathbb{P}_x(t < T^N)} \leq c_N(K);$$

(iv) $\lim_{N \to \infty} \|\kappa_N\|_\infty = 0$.

We denote by V a function satisfying Assumption 1.5 such that the H- exponents are positive. We use recent results in [CV17b] to prove that under the above conditions, Assumption 6.3 holds. More precisely, we have :

Proposition 6.3. *There exists $\theta > 0$ and $N_0 \geq 0$ such that, for all $N \geq N_0$, X^N admits a unique QSD α^N satisfying $\alpha^N(e^{\theta V}) < \infty$. Furthermore, there exist $\gamma_N \in (0, 1)$ and $C_N \geq 0$ such that, for all $t \geq 0$, for all $\mu \in \mathcal{P}(M_+)$,*

$$\|\mu \tilde{P}_t^N - \alpha^N\|_{TV} \leq C_N e^{-\gamma t} \mu(e^{\theta V}).$$

In addition, there is some $\delta > 0$ such that

$$\inf_{N \geq N_0} \alpha^N(M_{+,\delta}) > 0.$$

The last inequality implies that $\alpha(M_+) > 0$ for all limit point α of the sequence (α^N). Actually, we can prove that $\alpha(M_+) = 1$, and thus :

Proposition 6.4. *Let Π denote the unique invariant probability measure of X such that $\Pi(M_+) = 1$. Then :*

$$\lim_{N \to \infty} \alpha^N = \Pi.$$

6.5 Proof of main results

6.5.1 Proof of Theorem 6.1

This proof is inspired by [Mar13, Theorem 3.3]. Let U be a open subspace of M_+. Then, for all $t \geq 0$ and all N, one has by QSD property,

$$\begin{aligned}\alpha^N(U)e^{-\lambda_N t} &= \mathbb{P}_{\alpha^N}(X_t^N \in U)\\ &= \int \mathbb{P}_x(X_t^N \in U)\mathrm{d}\alpha^N(x)\\ &\geq \left(\inf_{x\in U} \mathbb{P}_x(X_t^N \in U)\right)\alpha^N(U).\end{aligned}$$

Thus if $\alpha^N(U) > 0$, one gets

$$e^{-\lambda_N t} \geq \inf_{x\in U} \mathbb{P}_x(X_t^N \in U) \tag{6.5}$$

Let (V, H) be functions satisfying Assumption 1.5 such that X is H-persistent. For some $R > 0$ to be chosen later, set $U = \{x : V(x) < R + 1\}$. By Assumption 6.3, there exists some $R > 0$ such that $\alpha^N(U) > 0$ for N large enough, in particular inequality (6.5) holds. For all $x^N \in M_+^N$ and $x \in M_+$, one has

$$\mathbb{P}_{x^N}(X_t^N \in U^c) = \mathbb{P}_{(x,x^N)}\left(V(X_t^N) \geq R+1; V(X_t) > R\right) + \mathbb{P}_{(x,x^N)}\left(V(X_t^N) \geq R+1; V(X_t) \leq R\right).$$

Now by compactness of M and continuity of V, $V_R = \{x : V(x) \leq R\}$ is compact and $\overline{V^c_{R+1}} = \{x : V(x) \geq R+1\}$ is closed, so there exists $\delta > 0$ such that $d(V_R, \overline{V^c_{R+1}}) = \delta$. In particular, $V(X_t^N) \geq R+1$ and $V(X_t) \leq R$ implies that $d(X_t^N, X_t) \geq \delta$. Thus, for all $x \in U$, $x^N \in U \cap M^N$ and $\theta > 0$,

$$\begin{aligned}\mathbb{P}_{x^N}(X_t^N \in U^c) &\leq \mathbb{P}_x\left(V(X_t) > R\right) + \mathbb{P}_{(x,x^N)}\left(d(X_t^N, X_t) \geq \delta\right)\\ &\leq \frac{\mathbb{E}_x(e^{\theta V(X_t)})}{e^{\theta R}} + \mathbb{P}_{(x,x^N)}\left(d(X_t^N, X_t) \geq \delta\right).\end{aligned}$$

Since X is H-persistent, by Theorem 1.13, there exist $\theta, \gamma, K > 0$, such that for all t big enough, for all $x \in M_+$, $\mathbb{E}_x(e^{\theta V(X_t)}) \leq e^{-\gamma t}e^{\theta V(x)} + K$. Hence, for all $x^N \in U$ and $\beta > 0$,

$$\mathbb{P}_{x^N}(X_t^N \in U^c) \leq e^{-\gamma t}e^{\theta} + Ke^{-\theta R} + \max_{x\in M_+,\, d(x^N,x)\leq\beta} \mathbb{P}_{(x,x^N)}\left(d(X_t^N, X_t) \geq \delta\right).$$

From 6.5, we then deduce that

$$\lambda_N \leq \frac{1}{t}\log\left(\frac{1}{1 - e^{-\gamma t}e^{\theta} - Ke^{-\theta R} - \max_{(x,x^N)\in M\times M^N,\, d(x^N,x)\leq\beta} \mathbb{P}_{(x,x^N)}\left(d(X_t^N, X_t) \geq \delta\right)}\right)$$

For all $\varepsilon > 0$, there exists $t, R > 0$ such that $e^{-\gamma t}e^{\theta} + Ke^{-\theta R} \leq \varepsilon$ and by (6.1), there exists N_0 and $\beta > 0$ such that for all $N \geq N_0$, $\max_{(x,x^N)\in M\times M^N,\, d(x^N,x)\leq\beta} \mathbb{P}_{(x,x^N)}\left(d(X_t^N, X_t) \geq \delta\right) \leq \varepsilon$, which proves that λ_N converges to 0. By Lemma 6.3, α is an invariant probability for X. □

6.5.2 Proof of Proposition 6.2

Recall that $M = [0,1]^d \times E$. We set $M_0 = \{0\} \times E$. With the assumptions on the rates B^k and C^k, M_0 is absorbing for X^N and satisfy point **(ii)** and **(iii)** of hypothesis 6.2.

Now it is a standard result which is a consequence of Perron-Frobenius Theorem that an irreducible absorbed continuous time Markov Chain admits a unique QSD. One can even show that the conditional law of X_t^N with respect to $T_0^N > t$ converges exponentially fast in total variation toward the QSD (see e.g. [CV16]).

The limit process Z satisfy hypothesis 6.5 **(i)** because 0 is invariant for each vector field F^i so it is also invariant for X. Moreover, Proposition 1.9 implies the Feller property and thus point **(ii)**.

We now prove the convergence (6.1) of Z^N toward Z for the metric d defined on $[0,1]^d \times E$ by

$$d((x,i);(y,j)) = \mathbb{1}_{i\neq j} + \mathbb{1}_{i=j}\|x-y\|.$$

More precisely, we show that for all $N>0$ and $\delta, T, \alpha > 0$, for all $(x,i) \in M$, $(x^N, i) \in M^N$ with $\|x - x^N\| \leq (\alpha \wedge \delta)/2e^{-KT}$,

$$\mathbb{P}_{(x,i);(x^N,i)}\left[\max_{[0,T]} d(Z_t, Z_t^N) \geq \delta\right] \leq H_N^T(\delta) + H_N^T(\alpha) + L\alpha\frac{\lambda}{\lambda+\sigma^N}.$$

where K is a common Lipschitz constant for the F^i and for all $\beta, T > 0$,

$$H_N^T(\beta) = 2d\exp\left(-\frac{\delta^2 e^{-2KT} N}{16CT}\right),$$

where C is a constant depending on T and δ given in Lemma 6.5 below. Note that since

$$\lim_{N\to\infty}\frac{N_i}{N} = p_i,$$

we have

$$\gamma := \sup_{N\geq 0}\max_i\left(\frac{N_i}{N}\right) < +\infty.$$

By adapting the proof of [Ben99, Proposition 4.6] and [BW03, Lemma 1], we show the following lemma :

Lemma 6.5. *For all $\varepsilon > 0$, for all $T > 0$, there exists $C > 0$ such that, for all $(x_0, i) \in M$, for all $N \geq 0$,*

$$\mathbb{P}_{x_0,i}\left[\max_{[0,T]}\|X_t^N - x_0 - \int_0^t F^{I_s^N}(X_s^N)ds\| \geq \varepsilon\right] \leq 2de^{-\frac{\varepsilon^2 N}{4CT}}. \tag{6.6}$$

Moreover, one can choose $C = \gamma e^{W_0(\varepsilon/2T)}$, where W_0 is the principal determination of the Lambert function.

Proof Let $(x_0, i) \in M$ and $C' > 0$. We claim that for $\theta = (\theta_1, \dots, \theta_d)$ with $\theta_i \leq N_i \log C'$, the process

$$Y_t^N = e^{\langle\theta, X_t^N - x_0 - \int_0^t F^{I_s^N}(X_s^N)ds\rangle - t\frac{C}{N}\|\theta\|^2}$$

is a supermartingale, where $C = C' \max_i \frac{1}{p_i}$. Before proving this claim, let us show how it implies the expected result. By maximal inequality for supermartingale, for all $\beta > 0$, it holds that

$$\mathbb{P}_{x_0,i}\left[\max_{[0,T]}\langle\theta, X_t^N - x_0 - \int_0^t F^{I_s^N}(X_s^N)ds\rangle \geq \beta\right] = \mathbb{P}_{x_0,i}\left[\max_{[0,T]} Y_t^N \geq e^{\beta - t\frac{C}{N}\|\theta\|^2}\right]$$
$$\leq e^{T\frac{C}{N}\|\theta\|^2-\beta}\mathbb{E}_{x_0,i}(Y_0^N) = e^{T\frac{C}{N}\|\theta\|^2-\beta}.$$

Set $\beta = \frac{N\varepsilon^2}{2CT}$, let $1 \leq i \leq d$ and set $\theta = \beta/\varepsilon e_i$. Then, $\theta_i \leq N_i \log C'$ if and only if $C \log C' \geq \frac{p_i\varepsilon}{2T}$. For such values of C we deduce that

$$\mathbb{P}_{x_0,i}\left[\max_{[0,T]}\langle e, X_t^N - x_0 - \int_0^t F^{I_s^N}(X_s^N)ds\rangle \geq \varepsilon\right] \leq e^{-\frac{\varepsilon^2 N}{4CT}},$$

which gives the announced result.

Now we prove the claim. Let $(x_0, i) \in M$, $\theta > 0$, and define $f : M \to \mathbb{R}_+$ by $f(x,i) = e^{\langle\theta, x-x_0\rangle}$. By classical results on Markov Process (see e.g. [EK86, Chapter 4, lemma 3.2]), the process

$$M_t^N = f(Z_t^N)e^{-\int_0^t \frac{L^N f(Z_s^N)}{f(Z_s^N)}ds}$$

is a martingale. Define the function g as $g(u) = e^u - u - 1$. Then

$$\frac{L^N f(x,k)}{f(x,k)} = \langle F^k(x), \theta\rangle + \sum_{i=1}^d g(\langle\theta, \frac{e_i}{N_i}\rangle)N_i B_i^k + \sum_{i=1}^d g(\langle\theta, -\frac{e_i}{N_i}\rangle)N_i C_i^k.$$

Let $C' > 0$. For all $u \leq \log C'$, $g(u) \leq 1/2u^2C'$. In particular, if $\theta = (\theta_1, \ldots, \theta_d)$ with $\theta_i \leq N_i \log C$, then

$$\sum_{i=1}^d g(\langle\theta, \frac{e_i}{N_i}\rangle)N_i B_i^k + \sum_{i=1}^d g(\langle\theta, -\frac{e_i}{N_i}\rangle)N_i C_i^k \leq C'\left(\max_i \frac{1}{p_i}\right)\frac{1}{N}\|\theta\|^2,$$

and thus

$$\frac{L^N f(x,k)}{f(x,k)} - \langle F^k(x), \theta\rangle \leq \frac{C}{N}\|\theta\|^2, \tag{6.7}$$

where $C = C'\gamma$. Now with the above notations, we have for all $u, t \geq 0$,

$$\frac{Y_{t+u}^N}{Y_t^N} = \frac{M_{t+u}^N}{M_t^N}\exp\left(-\int_t^{t+u}\langle\theta, F^{I_s^N}(X_s^N)\rangle + \frac{C}{N}\|\theta\|^2 - \frac{L^N f(Z_s^N)}{f(Z_s^N)}ds\right).$$

This concludes the proof of the claim because M^N is a martingale and, by (6.7), $\frac{Y_{t+u}^N}{Y_t^N} \leq \frac{M_{t+s}^N}{M_t^N}$. □

To give an insight on how to prove Proposition 6.2, we first prove it in the particular case when the jump rates $a_{k,l}$ do not depend on x. In this case, for all N, $(I_t^N)_{t\geq 0}$ is a continuous time Markov chain with rate matrix $(a_{k,l})_{k,l}$. Thus, we can assume without loss of generality that for all N and all $t \geq 0$, $I_t^N = I_t$ almost surely. In particular, $d(Z_t^N, Z_t) = \|X_t^N - X_t\|$.

For all $t \geq 0$, for all $(x^N, i) \in M$ and $(x, i) \in M$, one has

$$X_t^N = x^N + X_t^N - x^N - \int_0^t F^{I_s}(X_s^N)ds + \int_0^t F^{I_s}(X_s^N)ds, \tag{6.8}$$

and

$$X_t = x + \int_0^t F^{I_s}(X_s)ds. \tag{6.9}$$

Let K be a Lipschitz constant common for all the F^i. Due to (6.8) and (6.9), it holds that

$$\|X_t^N - X_t\| \leq \|x^N - x\| + V_N(t) + K\int_0^t \|X_s^N - X_s\|ds, \tag{6.10}$$

where $V_N(t) = \|X_t^N - x^N - \int_0^t F^{I_s}(X_s^N)ds\|$. By Grönwall lemma, we deduce that for all $T > 0$ and all $t \leq T$,

$$\|X_t^N - X_t\| \leq \left(\|x^N - x\| + D_N(T)\right) e^{KT},$$

where $D_N(T) = \max_{[0,T]} V_N(t)$. Hence, if $\|x^N - x\| \leq \delta/2e^{-KT}$, one gets

$$\begin{aligned}\mathbb{P}_{(x,k),(x^N,k)}\left[\max_{[0,T]} d(Z_t, Z_t^N) \geq \delta\right] &= \mathbb{P}_{(x,k),(x^N,k)}\left[\max_{[0,T]} \|X_t - X_t^N\| \geq \delta\right] \\ &\leq \mathbb{P}_{(x^N,k)}\left[D_N(T) \geq \frac{\delta}{2}e^{-KT}\right] \\ &\leq 2d\exp\left(-\frac{\delta^2 e^{-2KT}N}{16CT}\right),\end{aligned}$$

where the last inequality is a consequence of (6.6) and $C = \gamma e^{W_0(\frac{\delta}{4Te^{KT}})}$.

We have now to deal with the case where the jump rates are non constant. A first observation is the following. When the processes I and I^N start at the same point $k \in E$, they remain equal at least until a first jump occurs. In particular, if T_1 denotes this first time of jump, then (6.10) is still true for all $t \leq T_1$. As a consequence, if $T_1 \leq T$, X_{T_1} and $X_{T_1}^N$ are close with high probability. This fact encourages us to couple the jumps of I and I^N in such a manner that when X and X^N are close, I and I^N jump at the same time to the same state with high probability. We give an explicit construction of such a coupling.

Choose $\lambda, \sigma^N > 0$ such that

$$\lambda > \max_{(x,i)\in M} \sum_{i\in E} a_{ij}(x)$$

and

$$\sigma^N > \max_{(x,k)\in M} \sum_i N_i\left[B_i^k(x) + C_i^k(x)\right].$$

For all $x \in M$, define the stochastic matrix $(Q(x,i,j))_{i,j}$ by

$$Q(x,i,j) = \begin{cases} \frac{a(x,i,j)}{\lambda} & \text{if } i \neq j \\ 1 - \sum_{k\neq i} Q(x,i,k) & \text{if } i = j. \end{cases} \tag{6.11}$$

For all $k \in E$, define the stochastic matrix P_k^N on $\{0, \frac{1}{N_1}, \dots, 1\} \times \dots \times \{0, \frac{1}{N_d}, \dots, 1\}$ by

$$P_k^N(x,y) = \begin{cases} \frac{N_i B_i^k(x)}{\sigma^N} & \text{if } y = x + \frac{e_i}{N_i} \\ \frac{N_i C_i^k(x)}{\sigma^N} & \text{if } y = x - \frac{e_i}{N_i} \\ 1 - \sum_{z \neq x} P_k^N(x,z) & \text{if } y = x \\ 0 & \text{otherwise.} \end{cases} \tag{6.12}$$

Let V^N be a Poisson Process with parameter $\lambda + \sigma^N$. Let $(S_n)_{n\geq 0}$ denote the sequence of jump times of V^N, $(R_n)_{n\geq 1}$ be an i.i.d. sequence of Bernoulli variables with parameter $\frac{\lambda}{\lambda+\gamma^N}$, and $(W_n)_{n\geq 0}$ an i.i.d. sequence with uniform law on $[0,1]$, with (S_n), (R_n), (W_n) mutually independent. Without lose of generality, we set $E = \{1, \dots, |E|\}$.For all $(x,i) \in M$ and all $j \in E$, define the interval $E_{ij}(x) = [q_{j-1}(x,i), q_j(x,i))$ where $q_0(x,i) = 0$ and $q_j(x,i) = q_{j-1}(x,i) + Q(x,i,j)$. Note that for all x, the intervals $E_{ij}(x)$ form a partition of $[0,1)$. Now the processes (X,I) and (X^N, I^N) are constructed together as follows. Start from $(x,i) \in M$ and $(x^N, k) \in M^N$. Then for all $t < S_1$, $I_t = i$, $I_t^N = k$, $X_t^N = x^N$ and $X_t = \Phi_t^i(x)$, the flow generated by F^i. The process X is continuous, so $X_{S_1} = \Phi_{S_1}^i(x)$. If $R_1 = 0$, then $X_{S_1}^N = y$ with probability $P_k^N(X_{S_1}^N, y)$, and $I_{S_1} = I_{S_1}^N = i$. If $R_1 = 1$, then $X_{S_1}^N = x^N$ and there exists a unique couple $(j,l) \in E^2$ such that $W_1 \in E_{ij}(X_{S_1})$ and $W_1 \in E_{kl}(X_{S_1}^N)$. Set $I_{S_1} = j$ and $I_{S_1}^N = l$. This procedure in then repeated with starting point (X_{S_1}, I_{S_1}) and $(X_{S_1}^N, I_{S_1}^N)$ until time S_2. It it not hard to check that the process (X_t, I_t, X_t^N, I_t^N) so constructed is a coupling of Z and Z^N. With this construction, we get the following lemma :

Lemma 6.6. *There exists a constant $L > 0$ such that for all $\alpha > 0$, for all $N, n \geq 1$, for all $(x,i) \in M$, $(x^N, k) \in M^N$,*

$$\mathbb{P}_{(x,i),(x^N,i)}\left[I_{S_n} \neq I_{S_n}^N; \|X_{S_{n-1}} - X_{S_{n-1}}^N\| < \alpha | I_{S_{n-1}} = I_{S_{n-1}}^N\right] \leq L\alpha \frac{\lambda}{\lambda + \sigma^N}. \tag{6.13}$$

Proof For readability, we drop the dependence in the initial conditions. With the above construction, we have

$$\begin{aligned} &\mathbb{P}\left[I_{S_n} \neq I_{S_n}^N; \|X_{S_n} - X_{S_n}^N\| < \alpha | I_{S_{n-1}} = I_{S_{n-1}}^N\right] \\ &= \mathbb{E}\left[\mathbb{1}_{\|X_{S_{n-1}} - X_{S_{n-1}}^N\| < \alpha} \mathbb{P}\left(I_{S_n} \neq I_{S_n}^N | I_{S_{n-1}} = I_{S_{n-1}}^N; X_{S_{n-1}}; X_{S_{n-1}}^N\right)\right] \\ &= \mathbb{E}\left[\mathbb{1}_{\|X_{S_{n-1}} - X_{S_{n-1}}^N\| < \alpha} \sum_{i \in E} \mathbb{1}_{i = I_{S_{n-1}} = I_{S_{n-1}}^N} \mathbb{P}_{(X_{S_{n-1}},i),(X_{S_{n-1}}^N,i)}\left(I_{S_1} \neq I_{S_1}^N\right)\right] \\ &= \mathbb{E}\left[\mathbb{1}_{\|X_{S_{n-1}} - X_{S_{n-1}}^N\| < \alpha} \sum_{i \in E} \mathbb{1}_{i = I_{S_{n-1}} = I_{S_{n-1}}^N} \delta_i(X_{S_{n-1}}, X_{S_{n-1}}^N)\right]; \end{aligned}$$

where, for all $i \in E$, $\delta_i(x,y) = \mathbb{P}_{(x,i),(y,i)}\left(I_{S_1} \neq I_{S_1}^N\right)$. By construction, we have $\delta_i(x,y) = \mathbb{P}(R_1 = 1)\mathbb{P}\left(W_1 \in \bigcup_{k \neq l} E_{ik}(x) \cap E_{il}(y)\right)$. Since Q is Lipschitz, the function $(x,y) \mapsto \mathbb{P}\left(W_1 \in \bigcup_{k \neq l} E_{ik}(x) \cap E_{il}(y)\right)$ is also Lipschitz, with some constant L. Moreover, δ_i is 0 at $x = y$ hence $\delta_i(x,y) \leq \frac{\lambda}{\lambda+\sigma^N} L\|x - y\|$. This concludes the proof. $\square$

For all $(x,i) \in M$, $(x^N,i) \in M^N$, we have

$$\mathbb{P}_{(x,i);(x^N,i)}\left[\max_{[0,T]} d(Z_t, Z_t^N) \geq \delta\right] = \mathbb{P}_{(x,i);(x^N,i)}\left[\max_{[0,T]} d(Z_t, Z_t^N) \geq \delta; \max_{[0,T]} \mathbb{1}_{I_t^N \neq I_t} = 0\right] + \mathbb{P}_{(x,i);(x^N,i)}\left[\max_{[0,T]} d(Z_t, Z_t^N) \geq \delta; \max_{[0,T]} \mathbb{1}_{I_t^N \neq I_t} = 1\right]$$

On the event $\{\max_{[0,T]} \mathbb{1}_{I_t^N \neq I_t} = 0\}$, the two jumps processes remain equals until time T. Thus, on this event, $d(Z_t, Z_t^N) = \|X_t - X_t^N\|$ and inequality (6.10) holds for all $t \leq T$. In particular, we can conclude as in the case where the jump rates do not depend on the position that

$$\mathbb{P}_{(x,k),(x^N,k)}\left[\max_{[0,T]} d(Z_t, Z_t^N) \geq \delta; \max_{[0,T]} \mathbb{1}_{I_t^N \neq I_t} = 0\right] \leq 2d \exp\left(-\frac{\delta^2 e^{-2KT} N}{16CT}\right), \tag{6.14}$$

provided $\|x - x^N\| \leq \delta/2e^{-KT}$. For convenience, we write $H_N^T(\delta) = 2d \exp\left(-\frac{\delta^2 e^{-2KT} N}{16CT}\right)$. We introduce the random variable Y defined by $Y = \inf\{n \geq 1 \;:\; I_{S_n} \neq I_{S_n}^N\}$. Then

$$\mathbb{P}_{(x,i);(x^N,i)}\left[\max_{[0,T]} d(Z_t, Z_t^N) \geq \delta; \max_{[0,T]} \mathbb{1}_{I_t^N \neq I_t} = 1\right] \leq \mathbb{P}_{(x,i);(x^N,i)}\left[\max_{[0,T]} \mathbb{1}_{I_t^N \neq I_t} = 1\right] \leq \sum_{m=1}^{+\infty} \mathbb{P}(Y = m | S_m \leq T)\mathbb{P}(S_m \leq T).$$

Now we prove that for all $m \geq 1$, for all $\alpha > 0$ and x, x^N such that $\|x - x^N\| \leq \alpha/2e^{-KT}$,

$$\mathbb{P}_{(x,i);(x^N,i)}\left[Y = m | S_m \leq T\right] \leq H_N^T(\alpha) + L\alpha\frac{\lambda}{\lambda + \sigma^N}.$$

Note that we can rewrite the event $\{Y = m\}$ as

$$\{Y = m\} = \{I_{S_m} \neq I_{S_m}^N; I_{S_k} = I_{S_k}^N, \quad \forall 1 \leq k \leq m-1\}.$$

Let $\alpha > 0$. On the one hand, we have

$$\mathbb{P}_{(x,i);(x^N,i)}\left[Y = m; \|X_{S_{m-1}} - X_{S_{m-1}}^N\| \geq \alpha | S_m \leq T\right] \leq H_N^T(\alpha).$$

Indeed, on the event $\{Y = m\}$, inequality (6.10) holds until time S_{m-1}, which conditioned to $S_m \leq T$ is inferior to T. On the other hand, Lemma 6.6 implies that

$$\mathbb{P}_{(x,i);(x^N,i)}\left[Y = m; \|X_{S_{m-1}} - X_{S_{m-1}}^N\| \leq \alpha | S_m \leq T\right] \leq L\alpha\frac{\lambda}{\lambda + \sigma^N}.$$

Putting all together, we conclude that for all $N > 0$, $\delta, T, \alpha > 0$, for all $(x,i) \in M$, $(x^N,i) \in M^N$ with $\|x - x^N\| \leq (\alpha \wedge \delta)/2e^{-KT}$,

$$\mathbb{P}_{(x,i);(x^N,i)}\left[\max_{[0,T]} d(Z_t, Z_t^N) \geq \delta\right] \leq H_N^T(\delta) + H_N^T(\alpha) + L\alpha\frac{\lambda}{\lambda + \sigma^N}.$$

This ends the proof. □

6.5.3 Proof of Propositions 6.3 and 6.4

Proof of Proposition 6.3

As announced, we use a recent result of Champagnat and Villemonais to prove Proposition 6.3. That is, we show that **Condition (E)** of Section 1.5 is satisfied by the Markov chain $(X^N_{nT})_{n\geq 0}$ for some $T > 0$ and all N large enough. Furthermore, we prove that the constants $\theta_1, \theta_2, n_1, c_1, c_2$, the set K and the functions φ_1, φ_2 can be chosen independently from N, provided N is large enough.

We start with a lemma on the process X.

Lemma 6.7. *There exist $\theta > 0$, $\rho < \bar\rho < \max(1, \frac{\rho+1}{2})$, $C > 0$ and $T > 0$ such that :*

1.
$$P_T W \leq \bar\rho W + C\mathbb{1}_S,$$
where $W = e^{\theta V}$ and
$$S = \{x \in M_+ \ : W(x) \leq \frac{C}{\bar\rho - \rho}\}.$$

2. *There is a probability measure ν on S such that, for some constant c and for all $x \in S$,*
$$P_T(x, \cdot) \geq c\nu(\cdot \cap S).$$

3. *There exists $\gamma > 0$ scuh that, for all $x \in S$, for all $k \geq 0$,*
$$\mathbb{P}_x(X_{kT} \in S) \geq \gamma.$$

Proof By Theorem 1.13, for T_0 large enough, there exist $\theta, \delta, \kappa > 0$ and $\rho < 1$ such that, for all $t \in [T_0, T_0 + 1]$,

$$P_t(e^{\theta V}) \leq \rho e^{\theta V} + \kappa \mathbb{1}_{M_+ \setminus M_0^\delta}.$$

Set $W = e^{\theta V}$ and $C = \frac{\kappa}{1-\rho}$. Then, for all $n \geq 1$, for all $t \in [T_0, T_0 + 1]$,

$$P_{nt} W \leq \rho^n W + C \leq \rho W + C.$$

Now let $\bar\rho$ be such that $\rho < \bar\rho < \max(1, \frac{\rho+1}{2})$ and set

$$S = \{x \in M_+ \ : W(x) \leq \frac{C}{\bar\rho - \rho}\}.$$

Then, for all $n \geq 1$, for all $t \in [T_0, T_0 + 1]$,

$$P_{nt} W \leq \bar\rho W + C\mathbb{1}_S. \tag{6.15}$$

Now as S is a compact set of M_+, Proposition 1.8 and Assumption 6.6 imply that S is a small set. In particular, there exist $T_S, c_S > 0$, and a probability measure ν on S such that, for all $m \geq 1$, for all $x \in S$,

$$P_{mT_S}(x, \cdot) \geq c_S^m \nu(\cdot \cap S). \tag{6.16}$$

Now let $t \in [T_0, T_0+1]$ be such that $\frac{t}{T_S} = \frac{m}{n} \in \mathbb{Q}$, and set $T = nt = mT_S$. Then Equations (6.15) and (6.16) yield the two first points. Furthermore, let x be in S and $k \geq 0$. Then

$$\begin{aligned}
\mathbb{P}_x(X_{kT} \notin S) &= \mathbb{P}_x(W(X_{kT}) \geq \frac{C}{\bar{\rho}-\rho}) \\
&= \mathbb{P}_x(W(X_{knt}) \geq \frac{C}{\bar{\rho}-\rho}) \\
&\leq \frac{\bar{\rho}W(x)+C}{C}(\bar{\rho}-\rho) = 2\bar{\rho}-\rho.
\end{aligned}$$

This conclude the proof by setting $\gamma = 1-(2\bar{\rho}-\rho)$, which is positive by assumption. □

As $\|\kappa_N\|$ goes to 0 as N goes to infinity, let N_0 be such that

$$\theta_2 := \inf_{N \geq N_0} e^{-\|\kappa_N\|T} > \theta_1,$$

where $\theta_1 = \bar{\rho}$ and T are given in Lemma 6.7. Let $N \geq N_0$ and set $\mathcal{P} = \tilde{P}_T^N$. We prove that $\mathcal{P}$ satisfies Conditions (E1) to (E4) with $\theta_1 < \theta_2$, $\varphi_1 = W$ and $K = S$ given above and $\varphi_2 = \mathbb{1}_{M_+}$. To do so, we use the following straightforward lemma :

Lemma 6.8. *For all nonnegative measurable functions $f : M_+ \to \mathbb{R}$, for all $x \in M_+$, $t \geq 0$, $N \geq 0$,*

$$\tilde{P}_t^N f(x) = \mathbb{E}_x\left(f(X_t)e^{-\int_0^T \kappa_N(X_s)ds}\right).$$

In particular,

$$\tilde{P}_t^N f(x) \leq P_t f(x)$$

and

$$\tilde{P}_t^N f(x) \leq e^{-\|\kappa_N t\|} P_t f(x).$$

For all $x \in M_+$, we have by Lemma 6.7 and 6.8

$$\begin{aligned}
\mathcal{P}W(x) &\leq P_T W(x) \\
&\leq \theta_1 W(x) + C\mathbb{1}_S(x),
\end{aligned}$$

which implies the first two points of Condition (E2). Furthermore, we have for all $x \in M_+$

$$\begin{aligned}
\mathcal{P}\varphi_2(x) &= \mathbb{P}_x(X_T^N \in M_+) \\
&\geq e^{-\|\kappa_N T\|} \\
&\geq \theta_2\varphi_2(x),
\end{aligned}$$

by definition of θ_2. This yields the two last points of condition (E2). Now, still by Lemma 6.7 and 6.8, for all $x \in S$ and all $k \geq 1$

$$\mathbb{P}_x(X_{kT}^N \in S) \geq \gamma e^{-\|\kappa_N\|kT} > 0.$$

This implies Condition (E4). Condition (E1) is proved in a similar fashion with $n_1 = 1$, since

$$\begin{aligned}
\mathbb{P}_x(X_T^N \in \cdot) &\geq e^{-\|\kappa_N T\|}\mathbb{P}_x(X_T \in \cdot) \\
&\geq c\theta_2\nu(\cdot \cap S).
\end{aligned}$$

Thus, Theorem 1.21 implies that for all $N \geq N_0$, the Markov chain $(X^N_{nT})_{n\geq 0}$ admits a QSD α^N, which is the unique one satisfying $\alpha^N W < +\infty$. Furthermore, there exist C_N and $\gamma_N \in (0,1)$ such that, for all $n \geq 0$, for all $\mu \in \mathcal{P}(M_+)$ with $\mu W < +\infty$,

$$\|\mathbb{P}_\mu(X^N_{nT} \in \cdot | nT < T_N) - \alpha^N\|_{TV} \leq C_N \gamma_N^n \mu W. \tag{6.17}$$

In addition,

$$\inf_{N\geq N_0} \alpha^N(M_{+,\delta}) > 0,$$

where $\delta > 0$ is such that $S \subset M_{+,\delta}$. It remains to show that the same behaviour holds for the process $(X^N_t)_{t\geq 0}$. By the proof of Proposition 8.2 in [Ben18], there exists some constant $c > 0$ such that, for all $x \in M_+$ and all $t \geq 0$;

$$\mathbb{E}_x(e^{\theta V(X_t)}) \leq e^{\theta V(x)} e^{ct}.$$

Let $\mu \in \mathcal{P}(M_+)$ be such that $\mu W < +\infty$. Then the above inequality implies that

$$\mu P_t W \leq e^{ct} \mu W,$$

and thus

$$\mu \tilde{P}^N_t W \leq e^{ct} \mu W.$$

In particular, $(\mu \tilde{P}^N_t)W$ is finite for every $t \geq 0$. Therefore, by Equation 6.17, for all $t \geq 0$

$$\|\mathbb{P}_{\mu \tilde{P}^N_t}(X^N_{nT} \in \cdot | nT < T_N) - \alpha^N\|_{TV} \leq C_N \gamma_N^n \mu W.$$

This leads the result by the Markov property.

Proof of Proposition 6.4

Recall that for all $N \geq N_0$, for all $x \in M_+$,

$$\tilde{P}_T W(x) \leq \theta_1 W(x) + C \mathbb{1}_S.$$

Integrating this inequality with respect to α^N, one find

$$e^{-\lambda_N T} \alpha^N W \leq \theta_1 \alpha^N W + C.$$

Thus, since $\alpha^N W < +\infty$ and $e^{-\lambda_N T} \leq e^{-\|\kappa_N\| T}$, and by definition of θ_2,

$$\alpha^N W \leq \frac{C}{\theta_2 - \theta_1}.$$

This implies that the sequence $(\alpha^N)_{N\geq 0}$ is tight in M_+ and thus that every weak-limit point of $(\alpha^N)_{N\geq 0}$ is in $\mathcal{P}(M_+)$. So if α is such a limit point, $\alpha(M_+) = 1$, and by Theorem 6.1, $\alpha \in \mathcal{P}_{inv}(X)$. Now since X admits an accessible Doeblin point, Theorem 1.13 implies that $\mathcal{P}_{inv}(X) \cap \mathcal{P}(M_+)$ is reduced to a single element denoted Π, hence $\alpha = \Pi$. $\square$

Bibliography

[ABG+14] R. Azaïs, J.-B. Bardet, A. Génadot, N. Krell, and P.-A. Zitt, *Piecewise deterministic Markov process—recent results*, ESAIM: Proceedings, vol. 44, EDP Sciences, 2014, pp. 276–290.

[AC97] L. Arnold and N. D. Cong, *On the simplicity of the Lyapunov spectrum of products of random matrices*, Ergodic Theory Dynam. Systems **17** (1997), no. 5, 1005–1025. MR 1477030

[AC12] J-P. Aubin and A. Cellina, *Differential inclusions: set-valued maps and viability theory*, vol. 264, Springer Science & Business Media, 2012.

[AGD94] L. Arnold, V. M. Gundlach, and L. Demetrius, *Evolutionary formalism for products of positive random matrices*, Ann. Appl. Probab. **4** (1994), no. 3, 859–901. MR 1284989

[AHS79] L. Arnold, W. Horsthemke, and J. W. Stucki, *The influence of external real and white noise on the Lotka-Volterra model*, Biometrical Journal **21** (1979), no. 5, 451–471.

[ARBMW14] 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

[Arn98] L. Arnold, *Random dynamical systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998. MR 1723992

[Bax91] 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

[BBCD+18] J. Bierkens, A. Bouchard-Côté, A. Doucet, A. B. Duncan, P. Fearnhead, T. Lienart, G. Roberts, and S. J. Vollmer, *Piecewise deterministic Markov processes for scalable Monte Carlo on restricted domains*, Statist. Probab. Lett. **136** (2018), 148–154. MR 3806858

[BBM09] M Balde, U Boscain, and P. Mason, *A note on stability conditions for planar switched systems*, Internat. J. Control **82** (2009), no. 10, 1882–1888. MR 2567235

[BCG+13] J.-B. Bardet, A. Christen, A. Guillin, F. Malrieu, and P.-A. Zitt, *Total variation estimates for the TCP process*, Electron. J. Probab. **18** (2013), no. 10, 21. MR 3035738

[BCL17] M. Benaïm, F. Colonius, and R. Lettau, *Supports of invariant measures for piecewise deterministic Markov processes*, Nonlinearity **30** (2017), no. 9, 3400–3418. MR 3694261

[BD17] J. Bierkens and A. Duncan, *Limit theorems for the zig-zag process*, Adv. in Appl. Probab. **49** (2017), no. 3, 791–825. MR 3694318

[BEM07] J. Blath, A. Etheridge, and M. Meredith, *Coexistence in locally regulated competing populations and survival of branching annihilating random walk*, Ann. Appl. Probab. **17** (2007), no. 5-6, 1474–1507. MR 2358631

[Ben98] 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

[Ben99] Michel Benaïm, *Dynamics of stochastic approximation algorithms*, Séminaire de Probabilités, XXXIII, Lecture Notes in Math., vol. 1709, Springer, Berlin, 1999, pp. 1–68. MR 1767993

[Ben18] M. Benaïm, *Stochastic persistence*, `https://arxiv.org/abs/1806.08450`.

[BH99] 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

[BH12] Y. Bakhtin and T. Hurth, *Invariant densities for dynamical systems with random switching*, Nonlinearity **25** (2012), no. 10, 2937–2952. MR 2979976

[BHLM18] Y. Bakhtin, T. Hurth, S. Lawley, and J. Mattingly, *Smooth invariant densities for random switching on the torus*, Nonlinearity **31** (2018), no. 4, 1331.

[BHS08] M. Benaïm, J. Hofbauer, and W. H. Sandholm, *Robust permanence and impermanence for stochastic replicator dynamics*, J. Biol. Dyn. **2** (2008), no. 2, 180–195. MR 2427526

[BHS18] M. Benaïm, T. Hurth, and E. Strickler, *A user-friendly condition for exponential ergodicity in randomly switched environments*, Electron. Commun. Probab. **23** (2018), Paper No. 44, 12. MR 3841405

[BL85] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, Progress in Probability and Statistics, vol. 8, Birkhäuser Boston, Inc., Boston, MA, 1985. MR 886674

[BL16] M. Benaïm and C. Lobry, *Lotka Volterra in fluctuating environment or "how switching between beneficial environments can make survival harder"*, Ann. Appl. Probab. **26** (2016), no. 6, 3754–3785.

[BLBMZ12] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt, *Quantitative ergodicity for some switched dynamical systems*, Electron. Commun. Probab. **17** (2012), no. 56, 14. MR 3005729

[BLBMZ14] ______, *On the stability of planar randomly switched systems*, Ann. Appl. Probab. **24** (2014), no. 1, 292–311. MR 3161648

[BLBMZ15] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt, *Qualitative properties of certain piecewise deterministic Markov processes*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 3, 1040–1075. MR 3365972

[Bou87] P. Bougerol, *Limit theorem for products of random matrices with Markovian dependence*, Proceedings of the 1st World Congress of the Bernoulli Society, Vol. 1 (Tashkent, 1986), VNU Sci. Press, Utrecht, 1987, pp. 767–770. MR 1092426

[Bou88] ______, *Comparaison des exposants de Lyapounov des processus markoviens multiplicatifs*, Ann. Inst. H. Poincaré Probab. Statist. **24** (1988), no. 4, 439–489. MR 978021

[BRZ17] J. Bierkens, G. Roberts, and P.-A. Zitt, *Ergodicity of the zigzag process*, arXiv preprint arXiv:1712.09875 (2017).

[BS09] M. Benaïm and S. Schreiber, *Persistence of structured populations in random environments*, Theoretical Population Biology **76** (2009), 19–34.

[BS19] M. Benaïm and E. Strickler, *Random switching between vector fields having a common zero*, Ann. Appl. Probab. **29** (2019), no. 1, 326–375.

[BW03] M. Benaïm and J. W. Weibull, *Deterministic approximation of stochastic evolution in games*, Econometrica **71** (2003), no. 3, 873–903. MR 1983230

[CCM16] J.-R. Chazottes, P. Collet, and S. Méléard, *Sharp asymptotics for the quasi-stationary distribution of birth-and-death processes*, Probab. Theory Related Fields **164** (2016), no. 1-2, 285–332. MR 3449391

[CCM17] J-R Chazottes, P Collet, and S Méléard, *On time scales and quasi-stationary distributions for multitype birth-and-death processes*, arXiv preprint arXiv:1702.05369 (2017).

[CCPV18] N. Champagnat, K. A. Coulibaly-Pasquier, and D. Villemonais, *Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions*, Séminaire de Probabilités, XLIX, Lecture Notes in Math., vol. 2215, Springer, Berlin, 2018, pp. 165–182.

[CDG+17] B. Cloez, R. Dessalles, A. Genadot, F. Malrieu, A. Marguet, and R. Yvinec, *Probabilistic and piecewise deterministic models in biology*, ESAIM: Proceedings and Surveys **60** (2017), 225–245.

[CDMR12] A. Crudu, A. Debussche, A. Muller, and O. Radulescu, *Convergence of stochastic gene networks to hybrid piecewise deterministic processes*, Ann. Appl. Probab. **22** (2012), no. 5, 1822–1859. MR 3025682

[CH15] B. Cloez and M. Hairer, *Exponential ergodicity for Markov processes with random switching*, Bernoulli **21** (2015), 505–536.

[Che82] P. L. Chesson, *The stabilizing effect of a random environment*, Journal of Mathematical Biology **15** (1982), 1–36. MR 684776

[Che00] P. Chesson, *General theory of competitive coexistence in spatially-varying environments*, Theoretical Population Biology **58** (2000), no. 3, 211–237.

[Chu02] I. Chueshov, *Monotone random systems theory and applications*, Lecture Notes in Mathematics, vol. 1779, Springer-Verlag, Berlin, 2002. MR 1902500

[CM10] P. Cattiaux and S. Méléard, *Competitive or weak cooperative stochastic Lotka–Volterra systems conditioned on non-extinction*, J. Math. Biol. **60** (2010), no. 6, 797–829. MR 2606515

[CMP10] D. Chafaï, F. Malrieu, and K. Paroux, *On the long time behavior of the TCP window size process*, Stochastic Process. Appl. **120** (2010), no. 8, 1518–1534. MR 2653264

[Cra84] H. Crauel, *Lyapunov numbers of Markov solutions of linear stochastic systems*, Stochastics **14** (1984), no. 1, 11–28. MR 774582

[CV16] N. Champagnat and D. Villemonais, *Exponential convergence to quasi-stationary distribution and Q-process*, Probab. Theory Related Fields **164** (2016), no. 1-2, 243–283. MR 3449390

[CV17a] ———, *Exponential convergence to quasi-stationary distribution for absorbed one-dimensional diffusions with killing*, ALEA Lat. Am. J. Probab. Math. Stat. **14** (2017), no. 1, 177–199. MR 3622466

[CV17b] ———, *General criteria for the study of quasi-stationarity*, arXiv preprint arXiv:1712.08092 (2017).

[CV17c] ———, *Lyapunov criteria for uniform convergence of conditional distributions of absorbed markov processes*, arXiv preprint arXiv:1704.01928 (2017).

[CV17d] ———, *Uniform convergence to the Q-process*, Electron. Commun. Probab. **22** (2017), Paper No. 33, 7. MR 3663104

[CV18] ———, *Uniform convergence of conditional distributions for absorbed one-dimensional diffusions*, Adv. in Appl. Probab. **50** (2018), no. 1, 178–203. MR 3781982

[Dav84] 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. MR 790622

[Dav93] ———, *Markov models & optimization*, vol. 49, CRC Press, 1993.

[DGM18] A. Durmus, A. Guillin, and P. Monmarché, *Piecewise Deterministic Markov Processes and their invariant measure*, arXiv preprint arXiv:1807.05421 (2018).

[DMPS18] R. Douc, E. Moulines, P. Priouret, and P. Soulier, *Markov Chains*, first ed., Springer Series in Operations Research and Financial Engineering, Springer International Publishing, 2018.

[Duf97] M. Duflo, *Random iterative models*, Applications of Mathematics (New York), vol. 34, Springer-Verlag, Berlin, 1997, Translated from the 1990 French original by Stephen S. Wilson and revised by the author. MR 1485774

[EHS15] S. N. Evans, A. Hening, and S. J. Schreiber, *Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments*, J. Math. Biol. **71** (2015), no. 2, 325–359. MR 3367678

[EK86] S. N. Ethier and T. G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986, Characterization and convergence. MR 838085

[ERSS13] S. N. Evans, P. L. Ralph, S. J. Schreiber, and A. Sen, *Stochastic population growth in spatially heterogeneous environments*, J. Math. Biol. **66** (2013), no. 3, 423–476. MR 3010201

[FBPR18] P. Fearnhead, J. Bierkens, M. Pollock, and G. O. Roberts, *Piecewise deterministic Markov processes for continuous-time Monte Carlo*, Statist. Sci. **33** (2018), no. 3, 386–412. MR 3843382

[FGM12] J. Fontbona, H. Guérin, and F. Malrieu, *Quantitative estimates for the long-time behavior of an ergodic variant of the telegraph process*, Adv. in Appl. Probab. **44** (2012), no. 4, 977–994. MR 3052846

[FGM16] ———, *Long time behavior of telegraph processes under convex potentials*, Stochastic Process. Appl. **126** (2016), no. 10, 3077–3101. MR 3542627

[FK60] H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Statist. **31** (1960), 457–469. MR 0121828

[FMC09] 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

[FS14] M. Faure and S. J. Schreiber, *Quasi-stationary distributions for randomly perturbed dynamical systems*, Ann. Appl. Probab. **24** (2014), no. 2, 553–598. MR 3178491

[FW12] 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

[GH79] T. C. Gard and T. G. Hallam, *Persistence in food webs. I. Lotka-Volterra food chains*, Bull. Math. Biol. **41** (1979), no. 6, 877–891. MR 640001

[GH03] B. M. Garay and J. Hofbauer, *Robust permanence for ecological differential equations, minimax, and discretizations*, SIAM J. Math. Anal. **34** (2003), no. 5, 1007–1039. MR 2001657

[Gil75] M. E. Gilpin, *Group selection in predator-prey communities*, vol. 9, Princeton University Press, 1975.

[GMO08] L. Gerencsér, G. Michaletzky, and Z. Orlovits, *Stability of block-triangular stationary random matrices*, Systems Control Lett. **57** (2008), no. 8, 620–625. MR 2428831

[GSM07] L. Gurvits, R. Shorten, and O. Mason, *On the stability of switched positive linear systems*, IEEE Trans. Automat. Control **52** (2007), no. 6, 1099–1103. MR 2329904

[Has60] 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

[Hen84] H. Hennion, *Loi des grands nombres et perturbations pour des produits réductibles de matrices aléatoires indépendantes*, Z. Wahrsch. Verw. Gebiete **67** (1984), no. 3, 265–278. MR 762081

[Hir94] 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

[HM11] M. Hairer and J. C. Mattingly, *Yet another look at Harris' ergodic theorem for Markov chains*, Seminar on Stochastic Analysis, Random Fields and Applications VI, Progr. Probab., vol. 63, Birkhäuser/Springer Basel AG, Basel, 2011, pp. 109–117. MR 2857021

[HMS11] M. Hairer, J. C. Mattingly, and M. Scheutzow, *Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations*, Probab. Theory Related Fields **149** (2011), no. 1-2, 223–259. MR 2773030

[HN] A. Hening and D. Nguyen, *Coexistence and extinction for stochastic Kolmogorov systems*, Ann. Appl. Probab. **28**, no. 3, 1893–1942.

[HN18a] ———, *Persistence in stochastic Lotka-Volterra food chains with intraspecific competition*, Bulletin of Mathematical Biology (2018).

[HN18b] ———, *Stochastic Lotka-Volterra food chains*, Journal of Mathematical Biology **77** (2018), no. 1, 135–163.

[HNY18a] A. Hening, D. Nguyen, and G. Yin, *Stochastic population growth in spatially heterogeneous environments: The density-dependent case*, Journal of Mathematical Biology **76** (2018), no. 3, 697–754.

[HNY18b] A. Hening, D. H. Nguyen, and G. Yin, *Stochastic population growth in spatially heterogeneous environments: the density-dependent case*, J. Math. Biol. **76** (2018), no. 3, 697–754. MR 3747512

[Hof81] Josef Hofbauer, *A general cooperation theorem for hypercycles*, Monatshefte für Mathematik **91** (1981), no. 3, 233–240.

[HS74] MW Hirsch and S Smale, *Differential equations, dynamical systems, and linear algebra academic*, New York (1974), 2.

[HS98] J. Hofbauer and K. Sigmund, *Evolutionary games and population dynamics*, Cambridge university press, 1998.

[HS04] J. Hofbauer and S. J. Schreiber, *To persist or not to persist?*, Nonlinearity **17** (2004), no. 4, 1393–1406. MR 2069711

[HS19] A. Hening and E. Strickler, *On a predator-prey system with random switching that never converges to its equilibrium*, SIAM Journal on Mathematical Analysis (2019), accepted modulo revision.

[KK01] R. Z. Khasminskii and F. C. Klebaner, *Long term behavior of solutions of the lotka-volterra system under small random perturbations*, Ann. Appl. Probab. (2001), 952–963.

[Kur81] Thomas G. Kurtz, *Approximation of population processes*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 36, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1981. MR 610982

[Lag16] G. Lagasquie, *A note on simple randomly switched linear systems*, arXiv preprint arXiv:1612.01861 (2016).

[LES03] R. Lande, S. Engen, and B.-E. Saether, *Stochastic population dynamics in ecology and conservation*, Oxford University Press on Demand, 2003.

[LLC17] D. Li, S. Liu, and J. Cui, *Threshold dynamics and ergodicity of an SIRS epidemic model with Markovian switching*, J. Differential Equations **263** (2017), no. 12, 8873–8915. MR 3710707

[LMR14] S. D. Lawley, J. C. Mattingly, and M.l C. Reed, *Sensitivity to switching rates in stochastically switched ODEs*, Commun. Math. Sci. **12** (2014), no. 7, 1343–1352. MR 3210750

[Lot25] A. J. Lotka, *Elements of physical biology.*

[LS79] P. A. W. Lewis and G. S. Shedler, *Simulation of nonhomogeneous Poisson processes by thinning*, Naval Res. Logist. Quart. **26** (1979), no. 3, 403–413. MR 546120

[LY76] A. Lajmanovich and J. A. Yorke, *A deterministic model for gonorrhea in a nonhomogeneous population*, Math. Biosci. **28** (1976), no. 3/4, 221–236. MR 0403726

[Mar13] B. Marmet, *Quasi-stationary distributions for stochastic approximation algorithms with constant step size*, arXiv preprint arXiv:1303.7081 (2013).

[Mar14] ———, *Long time behaviour of 1/2 h\" older diffusion population processes*, arXiv preprint arXiv:1411.0533 (2014).

[MHP14] N. Massarelli, K. Hoffman, and J. P. Previte, *Effect of parity on productivity and sustainability of Lotka-Volterra food chains: bounded orbits in food chains*, J. Math. Biol. **69** (2014), no. 6-7, 1609–1626. MR 3275207

[Mie15] 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

[Mon16] P. Monmarché, *Piecewise deterministic simulated annealing*, ALEA Lat. Am. J. Probab. Math. Stat. **13** (2016), no. 1, 357–398. MR 3487077

[MP16] F. Malrieu and T. H. Phu, *Lotka-Volterra with randomly fluctuating environments: a full description*, arXiv preprint arXiv:1607.04395 (2016).

[MT09] S. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*, second ed., Cambridge University Press, Cambridge, 2009, With a prologue by Peter W. Glynn. MR 2509253

[MV12] S. Méléard and D. Villemonais, *Quasi-stationary distributions and population processes*, Probab. Surv. **9** (2012), 340–410. MR 2994898

[MZ17] F. Malrieu and P.-A. Zitt, *On the persistence regime for Lotka-Volterra in randomly fluctuating environments*, ALEA Lat. Am. J. Probab. Math. Stat. **14** (2017), no. 2, 733–749. MR 3695481

[Nua06] D. Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR 2200233

[Ose68] V. I. Oseledets, *A multiplicative ergodic theorem. characteristic Ljapunov, exponents of dynamical systems*, Trudy Moskovskogo Matematicheskogo Obshchestva **19** (1968), 179–210.

[PTW10a] K. Pakdaman, M. Thieullen, and G. Wainrib, *Fluid limit theorems for stochastic hybrid systems with application to neuron models*, Adv. in Appl. Probab. **42** (2010), no. 3, 761–794. MR 2779558

[PTW10b] K. Pakdaman, M. Thieullen, and G Wainrib, *Intrinsic variability of latency to first-spike*, Biol. Cybernet. **103** (2010), no. 1, 43–56. MR 2658681

[PTW12] ———, *Asymptotic expansion and central limit theorem for multiscale piecewise-deterministic Markov processes*, Stochastic Process. Appl. **122** (2012), no. 6, 2292–2318. MR 2922629

[RS14] G. Roth and S. J. Schreiber, *Persistence in fluctuating environments for interacting structured populations*, J. Math. Biol. **69** (2014), no. 5, 1267–1317. MR 3268345

[Rud03] R. Rudnicki, *Long-time behaviour of a stochastic prey-predator model*, Stochastic Process. Appl. **108** (2003), no. 1, 93–107. MR 2008602

[SBA11] S.n J. Schreiber, M.l Benaïm, and K. A. S. Atchadé, *Persistence in fluctuating environments*, J. Math. Biol. **62** (2011), no. 5, 655–683. MR 2786721

[Sch98] R. L. Schilling, *Conservativeness and extensions of Feller semigroups*, Positivity **2** (1998), no. 3, 239–256. MR 1653474

[Sch00] S. J. Schreiber, *Criteria for C^r robust permanence*, J. Differential Equations **162** (2000), no. 2, 400–426. MR 1751711

[Sch12] S. Schreiber, *Persistence for stochastic difference equations: A mini review*, Journal of Difference Equations and Applications **18** (2012), 1381–1403. MR 2956051

[Sen06] 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

[SJ72] Héctor J. Sussmann and Velimir Jurdjevic, *Controllability of nonlinear systems*, J. Differential Equations **12** (1972), 95–116.

[SLS09] S. J. Schreiber and J. O. Lloyd-Smith, *Invasion dynamics in spatially heterogeneous environments*, The American Naturalist **174** (2009), no. 4, 490–505.

[Str18] E. Strickler, *Randomly switched vector fields sharing a zero on a common invariant face*, arXiv preprint arXiv:1810.06331 (2018).

[TDHS06] Y. Takeuchi, N. H. Du, N. T. Hieu, and K. Sato, *Evolution of predator-prey systems described by a Lotka-Volterra equation under random environment*, J. Math. Anal. Appl. **323** (2006), no. 2, 938–957. MR 2260154

[Tuc99] W. Tucker, *The Lorenz attractor exists*, C. R. Acad. Sci. Paris Sér. I Math. **328** (1999), no. 12, 1197–1202. MR 1701385

[Vol28] V. Volterra, *Variations and fluctuations of the number of individuals in animal species living together*, J. Cons. Int. Explor. Mer **3** (1928), no. 1, 3–51.