

Qualitative properties of certain piecewise deterministic Markov processes

Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, Pierre-André Zitt

▶ **To cite this version:**

Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, Pierre-André Zitt. Qualitative properties of certain piecewise deterministic Markov processes. Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, 2015, 51 (3), pp.1040-1075. 10.1214/14-AIHP619 . hal-00688920v4

HAL Id: hal-00688920

https://hal.science/hal-00688920v4

Submitted on 7 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Qualitative properties of certain piecewise deterministic Markov processes

Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, Pierre-André Zitt

April 7, 2014

Abstract

We study a class of Piecewise Deterministic Markov Processes with state space $\mathbb{R}^d \times E$ where E is a finite set. The continuous component evolves according to a smooth vector field that is switched at the jump times of the discrete coordinate. The jump rates may depend on the whole position of the process. Working under the general assumption that the process stays in a compact set, we detail a possible construction of the process and characterize its support, in terms of the solutions set of a differential inclusion. We establish results on the long time behaviour of the process, in relation to a certain set of accessible points, which is shown to be strongly linked to the support of invariant measures. Under Hörmander-type bracket conditions, we prove that there exists a unique invariant measure and that the processes converges to equilibrium in total variation. Finally we give examples where the bracket condition does not hold, and where there may be one or many invariant measures, depending on the jump rates between the flows.

Keywords: Piecewise deterministic Markov Process, convergence to equilibrium, differential inclusion, Hörmander bracket condition

AMS Classification: 60J99, 34A60

1 Introduction

Piecewise deterministic Markov processes (PDMPs in short) are intensively used in many applied areas (molecular biology [27], storage modelling [6], Internet traffic [14, 17, 18], neuronal activity [7, 25],...). Roughly speaking, a Markov process is a PDMP if its randomness is only given by the jump mechanism: in particular, it admits no diffusive dynamics. This huge class of processes has been introduced by Davis [10]. See [11, 19] for a general presentation.

In the present paper, we deal with an interesting subclass of the PDMPs that plays a role in molecular biology [27, 7] (see also [29] for other motivations). We consider a PDMP evolving on $\mathbb{R}^d \times E$, where $d \geqslant 1$ and E is a finite set, as follows: the first coordinate moves continuously on $\mathbb{R}^d$ according to a smooth vector field that depends on the second coordinate, whereas the second coordinate jumps with a rate depending on the first one. Of course, most of the results in the present paper should extend to smooth manifolds. This class of Markov processes is reminiscent of the so-called iterated random functions in the discrete time setting (see [12] for a good review of this topic).

We are interested in the long time qualitative behaviour of these processes. A recent paper by Bakhtin and Hurth [2] considers the particular situation where the jump rates are constant and prove the beautiful result that, under a Hörmander type condition, if there exists an invariant measure for the process, then it is unique and absolutely continuous

with respect to the "Lebesgue" measure on $\mathbb{R}^d \times E$. Here we consider a more general situation and focus also on the convergence to equilibrium. We also provide a basic proof of the main result in [2].

Let us define our process more precisely. Let E be a finite set, and for any $i \in E$, $F^i : \mathbb{R}^d \mapsto \mathbb{R}^d$ be a smooth vector field. We assume throughout that each F^i is bounded and we denote by C_{sp} an upper bound for the "speed" of the deterministic dynamics:

$$\sup_{x \in \mathbb{R}^d, i \in E} \|F^i(x)\| \leqslant C_{sp} < \infty.$$

We let $\Phi^i = \{\Phi^i_t\}$ denote the flow induced by F^i. Recall that

$$t \mapsto \Phi^i_t(x) = \Phi^i(t, x)$$

is the solution to the Cauchy problem $\dot{x} = F^i(x)$ with initial condition $x(0) = x$. Moreover, we assume that there exists a compact set $M \subset \mathbb{R}^d$ that is positively invariant under each Φ^i, meaning that:

$$\forall i \in E,\ \forall t \geqslant 0, \quad \Phi^i_t(M) \subset M. \tag{1}$$

We consider here a continuous time Markov process $(Z_t = (X_t, Y_t))$ living on $M \times E$ whose infinitesimal generator acts on functions

$$\begin{aligned} g : M \times E &\to \mathbb{R}, \\ (x, i) &\mapsto g(x, i) = g^i(x), \end{aligned}$$

smooth[1] in x, according to the formula

$$Lg(x, i) = \langle F^i(x), \nabla g^i(x) \rangle + \sum_{j \in E} \lambda(x, i, j)(g^j(x) - g^i(x)) \tag{2}$$

where

(i) $x \mapsto \lambda(x, i, j)$ is continuous;

(ii) $\lambda(x, i, j) \geqslant 0$ for $i \neq j$ and $\lambda(x, i, i) = 0$;

(iii) for each $x \in M$, the matrix $(\lambda(x, i, j))_{ij}$ is irreducible.

The process is explicitly constructed in Section 2 and some of its basic properties (dynamics, invariant and empirical occupation probabilities) are established. In Section 3.1 we describe (Theorem 3.4) the support of the law of the process in term of the solutions set of a differential inclusion induced by the collection $\{F^i : i \in E\}$. Section 3.2 introduces the *accessible set* which is a natural candidate to support invariant probabilities. We show (Proposition 3.9) that this set is compact, connected, strongly positively invariant and invariant under the differential inclusion induced by $\{F^i : i \in E\}$. Finally, we prove that, if the process has a unique invariant probability measure, its support is characterized in terms of the accessible set.

Section 4 contains the main results of the present paper. We begin by a slight improvement of the regularity results of [2]: under Hörmander-like bracket conditions, the law of the process after a large enough number of jumps or at a sufficiently large time has an absolutely continuous component with respect to the Lebesgue measure on $\mathbb{R}^d \times E$. Moreover, this component may be chosen uniformly with respect to the initial distribution.

[1] meaning that g^i is the restriction to M of a smooth function on $\mathbb{R}^d$.

The proofs of these results are postponed to Sections 6 and 7. We use these estimates in Section 4.2 to establish the exponential ergodicity (in the sense of the total variation distance) of the process under study. In Section 5, we show that our assumptions are sharp thanks to several examples. In particular, we stress that, when the Hörmander condition is violated, the uniqueness of the invariant measure may depend on the jump mechanism between flows, and not only on the flows themselves.

Remark 1.1 (Quantitative results)**.** *In the present paper, we essentially deal with qualitative properties of the asymptotic behavior for a large class of PDMPs. Under more stringent assumptions, [4] gives an explicit rate of convergence in Wasserstein distance, via a coupling argument.*

Remark 1.2 (Compact state space)**.** *The main results in the present paper are still valid even if the state space is no longer compact provided that the excursions out of some compact sets are suitably controlled (say with a Lyapunov function). Nevertheless, as shown in [5], the stability of Markov processes driven by an infinitesimal generator as* (2) *may depend on the jump rates. As a consequence, it is difficult to establish, in our general framework, sufficient conditions for the stability of the process under study without the invariance assumption* (1)*; results in this direction may however be found in the recent [8].*

2 Construction and basic properties

In this section we explain how to construct explicitly the process $(Z_t)_{t\geqslant 0}$ driven by (2). Standard references for the construction and properties of more general PDMPs are the monographs [11] and [19]. In our case the compactness allows a nice construction via a discrete process whose jump times follow an homogeneous Poisson process, similar to the classical "thinning" method for simulating non-homogeneous Poisson processes (see [22, 28]).

2.1 Construction

Since M is compact and the maps $\lambda(\cdot, i, j)$ are continuous, there exists $\lambda \in \mathbb{R}_+$ such that

$$\max_{x\in M, i\in E} \sum_{j\in E, j\neq i} \lambda(x, i, j) < \lambda.$$

Let us fix such a λ, and let

$$Q(x, i, j) = \frac{\lambda(x, i, j)}{\lambda}, \text{ for } i \neq j \qquad \text{and} \qquad Q(x, i, i) = 1 - \sum_{j\neq i} Q(x, i, j) > 0.$$

Note that $Q(x) = (Q(x, i, j))_{i,j\in E}$ is an irreducible aperiodic Markov transition matrix and that (2) can be rewritten as

$$Lg = Ag + \lambda(Qg - g)$$

where

$$Ag(x, i) = \langle F^i(x), \nabla g^i(x)\rangle \quad \text{and} \quad Qg(x, i) = \sum_{j\in E} Q(x, i, j)g^j(x). \tag{3}$$

Let us first construct a discrete time Markov chain $(\tilde{Z}_n)_{n\geqslant 0}$. Let $(N_t)_{t\geqslant 0}$ be a homogeneous Poisson with intensity λ; denote by $(T_n)_{n\geqslant 0}$ its jump times and $(U_n)_{n\geqslant 0}$ its interarrival times.

Let $\tilde{Z}_0 \in M \times E$ be a random variable independent of $(N_t)_{t\geqslant 0}$. Define $(\tilde{Z}_n)_n = (\tilde{X}_n, \tilde{Y}_n)_n$ on $M \times E$ recursively by:

$$\tilde{X}_{n+1} = \Phi^{\tilde{Y}_n}\Big(U_{n+1}, \tilde{X}_n\Big),$$
$$\mathbb{P}\left[\tilde{Y}_{n+1} = j \Big| \tilde{X}_{n+1}, \tilde{Y}_n = i\right] = Q\Big(\tilde{X}_{n+1}, i, j\Big).$$

Now define $(Z_t)_{t\geqslant 0}$ via interpolation by setting

$$\forall t \in [T_n, T_{n+1}), \quad Z_t = \Big(\Phi^{\tilde{Y}_n}\Big(t - T_n, \tilde{X}_n\Big), \tilde{Y}_n\Big). \tag{4}$$

The memoryless property of exponential random variables makes $(Z_t)_{t\geqslant 0}$ a continuous time càdlàg Markov process. We let $P = (P_t)_{t\geqslant 0}$ denote the semigroup induced by $(Z_t)_{t\geqslant 0}$. Denoting by $\mathcal{C}_0$ (resp. $\mathcal{C}_1$) the set of real valued functions $f : M \times E \to \mathbb{R}$ that are continuous (resp continuously differentiable) in the first variable, we have the following result.

Proposition 2.1. *The infinitesimal generator of the semigroup $P = (P_t)_{t\geqslant 0}$ is the operator L given by* (2)*. Moreover, P_t is* Feller*, meaning that it maps $\mathcal{C}_0$ into itself and, for $f \in \mathcal{C}_0$, $\lim_{t\to 0} \|P_t f - f\| = 0$.*

The transition operator $\tilde{P}$ of the Markov chain $\tilde{Z}$ also maps $\mathcal{C}_0$ to itself, and if K_t and $\tilde{K}$ are defined by

$$K_t g(x,i) = g\Big(\Phi_t^i(x), i\Big) \quad \textit{and} \quad \tilde{K} f = \int_0^\infty \lambda e^{-\lambda t} K_t f \, dt, \tag{5}$$

then $\tilde{P}$ can be written as:

$$\begin{aligned}\tilde{P}g(x,i) &= \mathbb{E}\left[g(\tilde{Z}_1)\Big| Z_0 = (x,i)\right] = \int_0^\infty K_t Q g(x,i) \lambda e^{-\lambda t} dt \\ &= \tilde{K} Q g(x,i).\end{aligned} \tag{6}$$

Proof. For each $t \geqslant 0$, P_t acts on bounded measurable maps $g : M \times E \to \mathbb{R}$ according to the formula

$$P_t g(x,i) = \mathbb{E}\left[g(Z_t) | Z_0 = (x,i)\right].$$

For $t \geqslant 0$ let $J_t = K_t Q$. It follows from (4) that

$$P_t g = \sum_{n\geqslant 0} \mathbb{E}\left[\mathbb{1}_{\{N_t = n\}} J_{U_1} \circ \cdots \circ J_{U_n} \circ K_{t-T_n} g\right] \tag{7}$$

By Lebesgue continuity theorem and (7), $P_t g \in \mathcal{C}_0$ whenever $g \in \mathcal{C}_0$. Moreover, setting apart the first two terms in (7) leads to

$$P_t g = e^{-\lambda t} K_t g + \lambda e^{-\lambda t} \int_0^t K_u Q K_{t-u} g \, du + R(g,t) \tag{8}$$

where $|R(g,t)| \leqslant \|g\| \, \mathbb{P}\left[N_t > 1\right] = \|g\| \left(1 - e^{-\lambda t}(1+\lambda t)\right)$. Therefore $\lim_{t\to 0} \|P_t g - g\| = 0$.

The infinitesimal generator of (K_t) is the operator A defined by (3). Thus $\frac{1}{t}(K_t g - g) \to Ag$, therefore, by (8),

$$\frac{P_t g - g}{t} \xrightarrow[t\to 0]{} Ag - \lambda g + \lambda Q g,$$

and the result on (P_t) follows. The expression (6) of $\tilde{P}$ is a consequence of the definition of the chain. From (6) one can deduce that $\tilde{P}$ is also Feller. □

Remark 2.2 (Discrete chains and PDMPs)**.** *The chain $\tilde{Z}$ records all jumps of the discrete part Y of the PDMP Z, but, since $Q(x,i,i) > 0$, $\tilde{Z}$ also contains "phantom jumps" that cannot be seen directly on the trajectories of Z.*

Other slightly different discrete chains may crop up in the study of PDMPs. The most natural one is the process observed at (true) jump times. In another direction, the chain $(\Theta_n)_{n\in\mathbb{N}}$ introduced in [9] corresponds (in our setting) to the addition of phantom jumps at rate 1. For this chain (Θ_n), the authors prove (in a more general setting) equivalence between stability properties of the discrete and continuous time processes.

Similar equivalence properties will be shown below for our chain $\tilde{Z}$ (Proposition 2.4, Lemma 2.5). Its advantage lies in the simplicity of its definition; in particular it leads to a simulation method that does not require the integration of jump rates along trajectories.

Notation. Throughout the paper we may write $\mathbb{P}_{x,i}[\cdot]$ for $\mathbb{P}[\cdot|Z_0 = (x,i)]$ and $\mathbb{E}_{x,i}[\cdot]$ for $\mathbb{E}[\cdot|Z_0 = (x,i)]$.

2.2 First properties of the invariant probability measures

Let $\mathcal{M}(M\times E)$ (respectively $\mathcal{M}^+(M\times E)$ and $\mathcal{P}(M\times E)$) denote the set of signed (respectively positive, and probability) measures on $M\times E$. For $\mu\in\mathcal{M}(M\times E)$ and $f\in L^1(\mu)$ we write μf for $\int f d\mu$. Given a bounded operator $K:\mathcal{C}_0\to\mathcal{C}_0$ and $\mu\in\mathcal{M}(M\times E)$ we let $\mu K\in\mathcal{M}(M\times E)$ denote the measure defined by duality :

$$\forall g\in\mathcal{C}_0,\quad (\mu K)g = \mu(Kg).$$

The mappings $\mu\mapsto\mu P_t$ and $\mu\mapsto\mu\tilde{P}$ preserve the sets $\mathcal{M}^+(M\times E)$ and $\mathcal{P}(M\times E)$.

Definition 2.3 (Notation, stability)**.** *We denote by $\mathcal{P}_{inv}$ (resp. $\tilde{\mathcal{P}}_{inv}$) the set of invariant probabilities for (P_t) (resp. $\tilde{P}$),*

$$\begin{aligned}\mu\in\mathcal{P}_{inv} &\iff \forall t\geqslant 0,\quad \mu P_t = \mu;\\ \mu\in\tilde{\mathcal{P}}_{inv} &\iff \mu\tilde{P} = \mu.\end{aligned}$$

We say that the process Z (or $\tilde{Z}$) is stable *if it has a unique invariant probability measure.*

For $n\in\mathbb{N}^*$ and $t>0$ we let $\tilde{\Pi}_n$ and Π_t the (random) occupation measures defined by

$$\tilde{\Pi}_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{\tilde{Z}_k}\quad \textit{and}\quad \Pi_t = \frac{1}{t}\int_0^t \delta_{Z_t}.$$

By standard results for Feller chains on a compact space (see e.g. [13]), the set $\tilde{\mathcal{P}}_{inv}$ is non empty, compact (for the weak-$\star$ topology) and convex. Furthermore, with probability one every limit point of $(\Pi_n)_{n\geqslant 1}$ lies in $\tilde{\mathcal{P}}_{inv}$.

The following result gives an explicit correspondence between invariant measures for the discrete and continuous processes.

Proposition 2.4 (Correspondence for invariant measures)**.** *The mapping $\mu\mapsto\mu\tilde{K}$ maps $\tilde{\mathcal{P}}_{inv}$ homeomorphically onto $\mathcal{P}_{inv}$ and extremal points of $\tilde{\mathcal{P}}_{inv}$ (i.e. ergodic probabilities for $\tilde{P}$) onto extremal points of $\mathcal{P}_{inv}$ (ergodic probabilities for P_t).*

The inverse homeomorphism is the map $\mu\mapsto\mu Q$ restricted to $\mathcal{P}_{inv}$.

Consequently, the continuous time process (Z_t) is stable if and only if the Markov chain $(\tilde{Z}_n)$ is stable.

Proof. For all $f \in \mathcal{C}_1$, integrating by parts $\int_0^\infty \frac{dK_t f}{dt} e^{-\lambda t} dt$ and using the identities $\frac{dK_t f}{dt} = AK_t f = K_t A f$ leads to

$$\tilde{K}(\lambda I - A)f = \lambda f = (\lambda I - A)\tilde{K}f. \tag{9}$$

Let $\mu \in \mathcal{P}(M \times E)$. Then, using (9) and the form of L gives

$$\mu\tilde{K}Lf = \mu\tilde{K}(A - \lambda I)f + \lambda\mu\tilde{K}Qf = \lambda\Big(-\mu f + \mu\tilde{P}f\Big), \tag{10}$$

$$\mu L(\tilde{K}f) = \mu(A - \lambda I)\tilde{K}f + \lambda\mu Q\tilde{K}f = \lambda\Big(-\mu f + \mu Q\tilde{K}f\Big). \tag{11}$$

If $\mu \in \tilde{\mathcal{P}}_{inv}$, (10) implies $(\mu\tilde{K})Lf = 0$ for all $f \in \mathcal{C}_1$ and since $\mathcal{C}_1$ is dense in $\mathcal{C}_0$ this proves that $\mu\tilde{K} \in \mathcal{P}_{inv}$. Similarly, if $\mu \in \mathcal{P}_{inv}$, (11) implies $\mu = \mu Q\tilde{K}$. Hence $(\mu Q) = (\mu Q)\tilde{K}Q = (\mu Q)\tilde{P}$ proving that $\mu Q \in \tilde{\mathcal{P}}_{inv}$. Furthermore the identity $\mu = \mu Q\tilde{K}$ for all $\mu \in \mathcal{P}_{inv}$ shows that the maps $\mu \mapsto \mu\tilde{K}$ and $\mu \mapsto \mu Q$ are inverse homeomorphisms. □

Lemma 2.5 (Comparison of empirical measures)**.** *Let $f : M \times E \mapsto \mathbb{R}$ be a bounded measurable function. Then*

$$\lim_{t\to\infty} \Pi_t f - \tilde{\Pi}_{N_t}\tilde{K}f = 0 \quad \textit{and} \quad \lim_{n\to\infty} \Pi_{T_n} f - \tilde{\Pi}_n\tilde{K}f = 0$$

with probability one.

Proof. Decomposing the continuous time interval $[0, t]$ along the jumps yields:

$$\Pi_t f = \frac{N_t}{t}\left(\frac{1}{N_t}\sum_{i=0}^{N_t-1}\int_{T_i}^{T_{i+1}} f(Z_s)ds + r_t\right)$$

where $\|r_t\| \leqslant \|f\|\frac{U_{N_t+1}}{N_t}$.

Since $\lim_{t\to\infty}\frac{N_t}{t} = \lambda$ almost surely and $\mathbb{P}\left[U_n/n \geqslant \varepsilon\right] = e^{-\lambda n\varepsilon}$, $r_t \xrightarrow[t\to\infty]{a.s.} 0$, one has

$$\Pi_t f - \frac{1}{N_t}\sum_{i=0}^{N_t-1}\int_{T_i}^{T_{i+1}} f(Z_s)ds \xrightarrow[t\to\infty]{a.s.} 0.$$

Now, note that

$$\int_{T_i}^{T_{i+1}} f(Z_s)ds = \int_0^{U_{i+1}} f\Big(\phi_s^{\tilde{Y}^i}(\tilde{X}_i), \tilde{Y}_i\Big)ds,$$

therefore

$$M_n = \sum_{i=0}^{n-1}\left(\int_{T_i}^{T_{i+1}} f(Z_s)ds - \tilde{K}f(\tilde{X}_i, \tilde{Y}_i)\right)$$

is a martingale with increments bounded in L^2: $\mathbb{E}\left[(M_{n+1} - M_n)^2\right] \leqslant 2\|f\|^2/\lambda^2$. Therefore, by the strong law of large numbers for martingales, $\lim_{n\to\infty}\frac{M_n}{n} = 0$ almost surely, and the result follows. □

As in the discrete time framework, the set $\mathcal{P}_{inv}$ is non empty compact (for the weak-$\star$ topology) and convex. Furthermore, with probability one, every limit point of $(\Pi_t)_{t\geqslant 0}$ lies in $\mathcal{P}_{inv}$.

Finally, one can check that an invariant measure for the embedded chain and its associated invariant measure for the time continuous process have the same support. Given $\mu \in \mathcal{P}(M \times E)$ let us denote by $\mathrm{supp}(\mu)$ its support.

Lemma 2.6. *Let $\mu \in \tilde{\mathcal{P}}_{inv}$. Then μ and $\mu\tilde{K}$ have the same support.*

Proof. Let $(x,i) \in \mathrm{supp}(\mu)$ and let $\mathcal{U}$ be a neighborhood of x. Then for $t_0 > 0$ small enough and $0 \leqslant t \leqslant t_0$, $\Phi^i_{-t}(\mathcal{U})$ is also a neighborhood of x. Thus

$$(\mu\tilde{K})(\mathcal{U}\times\{i\}) = \int_0^\infty \lambda e^{-\lambda t}\mu(\Phi^i_{-t}(\mathcal{U})\times\{i\})\,dt \geqslant \lambda\int_0^{t_0} e^{-\lambda t}\mu(\Phi^i_{-t}(\mathcal{U})\times\{i\})\,dt > 0.$$

This proves that $\mathrm{supp}(\mu) \subset \mathrm{supp}(\mu\tilde{K})$. Conversely, let $\nu = \mu\tilde{K}$ and $(x,i) \in \mathrm{supp}(\nu)$ and let $\mathcal{U}$ be a neighborhood of x. Then

$$\mu(\mathcal{U}\times\{i\}) = (\nu Q)(\mathcal{U}\times\{i\}) = \sum_{j\in E}\int_{\mathcal{U}} Q_{ji}(x)\nu(dx\times\{j\}) \geqslant \int_{\mathcal{U}} Q_{ii}(x)\nu(dx\times\{i\}) > 0.$$

As a consequence, $\mathrm{supp}(\mu) \supset \mathrm{supp}(\mu\tilde{K})$. □

2.3 Law of pure types

Let λ_M and $\lambda_{M\times E}$ denote the Lebesgue measures on M and $M\times E$.

Proposition 2.7 (Law of pure types)**.** *Let $\mu \in \tilde{\mathcal{P}}_{inv}$ (respectively $\mathcal{P}_{inv}$) and let $\mu = \mu_{ac}+\mu_s$ be the Lebesgue decomposition of μ with μ_{ac} the absolutely continuous (with respect to $\lambda_{M\times E}$) measure and μ_s the singular (with respect to $\lambda_{M\times E}$) measure. Then both μ_{ac} and μ_s are in $\tilde{\mathcal{P}}_{inv}$ (respectively $\mathcal{P}_{inv}$), up to a multiplicative constant. In particular, if μ is ergodic, then μ is either absolutely continuous or singular.*

Proof. The key point is that $\tilde{K}$ and Q, hence $\tilde{P} = \tilde{K}Q$, map absolutely continuous measures into absolutely continuous measures. For $\mu \in \tilde{\mathcal{P}}_{inv}$ the result now follows from the following simple Lemma 2.8 applied to $\tilde{P}$. □

Lemma 2.8. *Let $(\Omega,\mathcal{A},P)$ be a probability space. Let $\mathcal{M}$ (respectively $\mathcal{M}^+,\mathcal{P},\mathcal{M}_{ac}$) denote the set of signed (positive, probability, absolutely continuous with respect to P) measures on Ω. Let $K : \mathcal{M}\to\mathcal{M}, \mu\mapsto\mu K$ be a linear map that maps each of the preceding sets into itself. Then if $\mu\in\mathcal{P}$ is a fixed point for K with Lebesgue decomposition $\mu = \mu_{ac}+\mu_s$, both μ_{ac} and μ_s are fixed points for K.*

Proof. Write $\mu K = \mu_{ac}K + \mu_s K = \mu_{ac}K + \nu_{ac} + \nu_s$ with $\mu_s K = \nu_{ac}+\nu_s$ the Lebesgue decomposition of $\mu_s K$. Then, by uniqueness of the decomposition, $\mu_{ac} = \mu_{ac}K + \nu_{ac}$. Thus, $\mu_{ac} \geqslant \mu_{ac}K$. Now either $\mu_{ac} = 0$ and there is nothing to prove or, we can normalize by $\mu_{ac}(\Omega)$ and we get that $\mu_{ac} = \mu_{ac}K$. □

3 Supports and accessibility

3.1 Support of the law of paths

In this section, we describe the shape of the support of the distribution of $(X_t)_{t\geqslant 0}$ and we show that it can be linked to the set of solutions of a differential inclusion (which is a generalisation of ordinary differential equations).

Let us start with a definition that will prove useful to encode the paths of the process Z_t.

Definition 3.1 (Trajectories and adapted sequences)**.** *For all $n\in\mathbb{N}^*$ let*

$$\mathbb{T}_n = \left\{(\mathbf{i},\mathbf{u}) = ((i_0,\dots,i_n),(u_1,\dots,u_n)) \in E^{n+1}\times\mathbb{R}_+^n\right\}$$

and

$$\mathbb{T}_n^{i,j} = \{(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n \ : \ i_0 = i,\, i_n = j\}.$$

Given $(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n$ *and* $x \in M$*, define* $(x_k)_{0\leqslant k\leqslant n}$ *by induction by setting* $x_0 = x$ *and* $x_{k+1} = \Phi_{u_k}^{i_{k-1}}(x_k)$*: these are the points obtained by following* F^{i_0} *for a time* u_0*, then* F^{i_1} *for a time* u_1*, etc.*

We also define a corresponding continuous trajectory. Let $\mathbf{t} = (t_0,\dots,t_n)$ *be defined by* $t_0 = 0$ *and* $t_k = t_{k-1} + u_k$ *for* $k = 1,2,\dots,n$ *and let* $(\eta_{x,\mathbf{i},\mathbf{u}}(t))_{t\geqslant 0}$ *be the function* $(\eta(t))_{t\geqslant 0}$ *given by*

$$\eta_{x,\mathbf{i},\mathbf{u}}(t) = \eta(t) = \begin{cases} x & \text{if } t = 0, \\ \Phi_{t-t_{k-1}}^{i_{k-1}}(x_{k-1}) & \text{if } t_{k-1} < t \leqslant t_k \text{ for } k = 1,\dots,n, \\ \Phi_{t-t_n}^{i_n}(x_n) & \text{if } t > t_n. \end{cases} \tag{12}$$

Finally, let $p(x,\mathbf{i},\mathbf{u}) = Q(x_1,i_0,i_1)Q(x_2,i_1,i_2)\cdots Q(x_n,i_{n-1},i_n)$*, and*

$$\mathbb{T}_{n,\mathrm{ad}(x)} = \{(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n \ : \ p(x,\mathbf{i},\mathbf{u}) > 0\}.$$

An element of $\mathbb{T}_{n,\mathrm{ad}(x)}$ *is said to be* adapted to $x \in M$*.*

Lemma 3.2. *Let* $(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n$*. Then, for any* $(x,i) \in M \times E$*, any* $T \geqslant 0$*, and any* $\delta > 0$*,*

$$\mathbb{P}_{x,i}\left[\sup_{0\leqslant t\leqslant T} \|X_t - \eta_{x,\mathbf{i},\mathbf{u}}(t)\| \leqslant \delta\right] > 0.$$

Proof. Suppose first that $(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n$ is adapted to x and that $\mathbf{i}$ starts at i. By continuity, there exist δ_1 and δ_2 such that

$$\max_{i=1,\dots,n} |s_i - u_i| \leqslant \delta_1 \quad \Longrightarrow \quad \begin{cases} \sup_{0\leqslant t\leqslant T} \|\eta_{x,\mathbf{i},\mathbf{s}}(t) - \eta_{x,\mathbf{i},\mathbf{u}}(t)\| \leqslant \delta, \\ p(x,\mathbf{i},\mathbf{s}) \geqslant \delta_2. \end{cases}$$

Let $(U_1,\dots,U_n,U_{n+1})$ be $n+1$ independent random variables with an exponential law of parameter λ and $\mathbf{U} = (U_1,\dots,U_n)$. Then,

$$\begin{aligned} &\mathbb{P}_{x,i}\left[\sup_{0\leqslant t\leqslant T} \|X_t - \eta_{x,\mathbf{i},\mathbf{u}}(t)\| \leqslant \delta\right] \\ &\quad \geqslant \mathbb{P}\left[\max_{l=1,\dots,n} |U_l - u_l| \leqslant \delta_1,\ U_{n+1} \geqslant T - t_n + \delta_1, (\tilde{Y}_0,\dots,\tilde{Y}_n) = \mathbf{i}\right] \\ &\quad \geqslant \delta_2 \mathbb{P}\left[\max_{l=1,\dots,n} |U_l - u_l| \leqslant \delta_1,\ U_{n+1} \geqslant T - t_n + \delta_1\right] \\ &\quad \geqslant \delta_2 \left[\prod_{l=1}^{n} \left(e^{-\lambda(u_l-\delta_1)} - e^{-\lambda(u_l+\delta_1)}\right)\right] e^{-\lambda(T-t_n+\delta_1)} > 0. \end{aligned}$$

In the general case, $(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n$ is not necessarily adapted and may start at an arbitrary i_0. However, for any $T > 0$ and $\delta > 0$, there exists $(\mathbf{j},\mathbf{v}) \in \mathbb{T}_{n',\mathrm{ad}(x)}$ (for some $n' \geqslant n$) such that $j_0 = i$ and

$$\sup_{0\leqslant t\leqslant T} \|\eta_{x,\mathbf{j},\mathbf{v}}(t) - \eta_{x,\mathbf{i},\mathbf{u}}(t)\| \leqslant \delta,$$

since $Q(x)$ is, by construction, irreducible and aperiodic (this allows to add permitted transitions from i to i_0 and where $\mathbf{i}$ has not permitted transitions, with times between the jumps as small as needed). □

After these useful observations, we can describe the support of the law of $(X_t)_{t\geqslant 0}$ in terms of a certain differential inclusion induced by the vector fields $\{F^i : i \in E\}$.

For each $x \in \mathbb{R}^d$, let $\mathrm{co}(F)(x) \subset \mathbb{R}^d$ be the compact convex set defined as

$$\mathrm{co}(F)(x) = \left\{ \sum_{i\in E} \alpha_i F^i(x) : \alpha_i \geqslant 0, \sum_{i\in E} \alpha_i = 1 \right\}.$$

Let $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ denote the set of continuous paths $\eta : \mathbb{R}_+ \mapsto \mathbb{R}^d$ equipped with the topology of uniform convergence on compact intervals. A solution to the differential inclusion

$$\dot{\eta} \in \mathrm{co}(F)(\eta) \tag{13}$$

is an absolutely continuous function $\eta \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ such that $\dot{\eta}(t) \in \mathrm{co}(F)(\eta(t))$ for almost all $t \in \mathbb{R}_+$. We let $S^x \subset \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ denote the set of solutions to (13) with initial condition $\eta(0) = x$.

Lemma 3.3. *The set S^x is a non empty compact connected set.*

Proof. Follows from standard results on differential inclusion, since the set-valued map $\mathrm{co}(F)$ is upper-semi continuous, bounded with non empty compact convex images; see [1] for details. □

Theorem 3.4. *If $X_0 = x \in M$ then, the support of the law of $(X_t)_{t\geqslant 0}$ equals S^x.*

Proof. Obviously, any path of X is a solution of the differential inclusion (13). Let $\eta \in S^x$, and $\varepsilon > 0$. Set

$$G_t(x) = \Big\{ v \in \Big\{ F^i(x),\ i \in E \Big\} : \langle v - \dot{\eta}(t), x - \eta(t) \rangle < \varepsilon \Big\}.$$

Since $\dot{\eta}(t) \in \mathrm{co}(F)(\eta(t))$ almost surely, $G_t(x)$ is non empty. Furthermore, $(t, x) \mapsto G_t(x)$ is uniformly bounded, lower semicontinuous in x, and measurable in t. Hence, using a result by Papageorgiou [26], there exists $\xi : \mathbb{R} \to \mathbb{R}^d$ absolutely continuous such that $\xi(0) = x$ and $\dot{\xi}(t) \in G_t(\xi(t))$ almost surely. In particular,

$$\frac{d}{dt}\|\xi(t) - \eta(t)\|^2 = 2\langle \dot{\xi}(t) - \dot{\eta}(t), \xi(t) - \eta(t) \rangle < 2\varepsilon$$

so that

$$\sup_{0\leqslant t\leqslant T} \|\xi(t) - \eta(t)\|^2 \leqslant 2\varepsilon T.$$

Thus, without loss of generality, one can assume that η is such that

$$\dot{\eta}(t) \in \bigcup_{i\in E} \Big\{ F^i(\eta(t)) \Big\}$$

for almost all $t \in \mathbb{R}_+$. Set

$$\forall i \in E, \qquad \Omega_i = \Big\{ t \in [0, T] \ : \dot{\eta}(t) = F^i(\eta(t)) \Big\}.$$

Let $\mathcal{C}$ be the algebra consisting of finite unions of intervals in $[0, T]$. Since the Borel σ-field over $[0, T]$ is generated by $\mathcal{C}$, there exists, for all $i \in E$, $J_i \in \mathcal{C}$ such that, the set $\{J_i : i \in E\}$ forms a partition of $[0, T]$, and for $i \in E$

$$\lambda(\Omega_i \Delta J_i) \leqslant \varepsilon$$

where λ stands for the Lebesgue measure over $[0,T]$ and $A\Delta B$ is the symmetric difference of A and B. Hence, there exist numbers $0 = t_0 < t_1 < \cdots < t_{N+1} = T$ and a map $i : k \mapsto i_k$ from $\{0,\dots,N\}$ to $\{1,\dots,n_0\}$, such that $(t_k, t_{k+1}) \subset J_{i_k}$. Introduce $\eta_{x,\mathbf{i},\mathbf{u}}$ given by Formula (12). For all $t_k \leqslant t \leqslant t_{k+1}$,

$$\begin{aligned}\eta_{x,\mathbf{i},\mathbf{u}}(t) - \eta(t) =& \eta_{x,\mathbf{i},\mathbf{u}}(t_k) - \eta(t_k) \\ &+ \int_{t_k}^{t} \Big(F^{i_k}(\eta_{x,\mathbf{i},\mathbf{u}}(s)) - F^{i_k}(\eta(s))\Big)\, ds + \int_{t_k}^{t} \Big(F^{i_k}(\eta(s)) - \dot\eta(s)\Big)\, ds.\end{aligned}$$

Hence, by Gronwall's lemma, we get that

$$\|\eta_{x,\mathbf{i},\mathbf{u}}(t) - \eta(t)\| \leqslant e^{K(t_{k+1}-t_k)}(\|\eta_{x,\mathbf{i},\mathbf{u}}(t_k) - \eta(t_k)\| + m_k)$$

where K is a Lipschitz constant for all the vector fields (F^i) and

$$m_k = 2C_{sp}\lambda([t_k, t_{k+1}] \setminus \Omega_{i_k}).$$

It then follows that, for all $k = 0,\dots,N$ and $t_k \leqslant t \leqslant t_{k+1}$,

$$\|\eta_{x,\mathbf{i},\mathbf{u}}(t) - \eta(t)\| \leqslant \sum_{l=0}^{k} e^{K(t_{k+1}-t_l)} m_l \leqslant e^{KT} \sum_{l=0}^{N} m_l \leqslant e^{KT} \sum_{i=1}^{|E|} \lambda(J_i \setminus \Omega_i) \leqslant e^{KT}\varepsilon.$$

This, with Lemma 3.2, shows that S^x is included in the support of the law of $(X_t)_{t\geqslant 0}$ and concludes the proof. □

In the course of the proof, one has obtained the following result which is stated separately since it will be useful in the sequel.

Lemma 3.5. *If $\eta :\ \mathbb{R}_+ \to \mathbb{R}^d$ is such that*

$$\dot\eta(t) \in \bigcup_{i\in E} \Big\{F^i(\eta(t))\Big\}$$

for almost all $t \in \mathbb{R}_+$, then, for any $\varepsilon > 0$ and any $T > 0$, there exists $(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n$ (for some n) such that

$$\|\eta_{x,\mathbf{i},\mathbf{u}}(t) - \eta(t)\| \leqslant \varepsilon.$$

3.2 The accessible set

In this section we define and study the *accessible set* of the process X as the set of points that can be "reached from everywhere" by X and show that long term behavior of X is related to this set.

Definition 3.6 (Positive orbit and accessible set)**.**

For $(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n$, let $\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}$ be the "composite flow":

$$\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x) = \Phi^{i_{n-1}}_{u_n} \circ \ldots \circ \Phi^{i_0}_{u_1}(x). \tag{14}$$

The positive *orbit of x is the set*

$$\gamma^+(x) = \left\{ \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x) \ : (\mathbf{i},\mathbf{u}) \in \bigcup_{n\in\mathbb{N}^*} \mathbb{T}_n \right\}.$$

The accessible set *of (X_t) is the (possibly empty) compact set $\Gamma \subset M$ defined as*

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

Remark 3.7. *If $y = \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x)$ for some $(\mathbf{i}, \mathbf{u})$ then y is the limit of points of the form $\mathbf{\Phi}^{\mathbf{j}}_{\mathbf{v}}(x)$ where $(\mathbf{j}, \mathbf{v})$ is adapted to x. It implies that*

$$\Gamma = \bigcap_{x \in M} \overline{\{\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x) \ : (\mathbf{i}, \mathbf{u}) \quad \textit{adapted to } x\}}. \tag{15}$$

Remark 3.8. *The accessible set Γ is called the set of D-approachable points and is denoted by L in [2].*

3.2.1 Topological properties of the accessible set

The differential inclusion (13) induces a set-valued dynamical system $\Psi = \{\Psi_t\}$ defined by

$$\Psi_t(x) = \Psi(t, x) = \{\eta(t) : \ \eta \in S^x\}$$

enjoying the following properties

(i) $\Psi_0(x) = \{x\}$,

(ii) $\Psi_{t+s}(x) = \Psi_t(\Psi_s(x))$ for all $t, s \geqslant 0$,

(iii) $y \in \Psi_t(x) \Rightarrow x \in \Psi_{-t}(y)$.

For subsets $I \subset \mathbb{R}$ and $A \subset \mathbb{R}^d$ we set

$$\Psi(I, A) = \bigcup_{(t,x) \in I \times A} \Psi_t(x).$$

A set $A \subset \mathbb{R}^d$ is called *strongly* positively invariant under Ψ if $\Psi_t(A) \subset A$ for $t \geqslant 0$. It is called *invariant* if for all $x \in A$ there exists $\eta \in S^x_{\mathbb{R}}$ such that $\eta(\mathbb{R}) \subset A$. Given $x \in \mathbb{R}^d$, the *(omega) limit set* of x under Ψ is defined as

$$\omega_\Psi(x) = \bigcap_{t \geqslant 0} \overline{\Psi_{[t,\infty[}(x)}.$$

Lemma 3.9. *The set $\omega_\Psi(x)$ is compact connected invariant and strongly positively invariant under Ψ.*

Proof. It is not hard to deduce the first three properties from Lemma 3.3. For the last one, let $p \in \omega_\Psi(x)$, $s > 0$ and $q \in \Psi_s(p)$. By Lemma 3.5, for all $\varepsilon > 0$, there exists $n \in \mathbb{N}$ and $(\mathbf{i}, \mathbf{u}) \in \mathbb{T}_n$ such that $\left\| \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(p) - q \right\| < \varepsilon$. Continuity of $\Phi^{\mathbf{i}}_{\mathbf{u}}$ makes the set

$$\mathcal{W}_\varepsilon = \{z \in M : \ \left\| \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(z) - q \right\| < \varepsilon\}$$

an open neighborhood of p. Hence $\mathcal{W}_\varepsilon \cap \Psi_{[t,\infty[}(x) \neq \emptyset$ for all $t > 0$. This proves that the distance between the sets $\{q\}$ and $\Psi_{[t,\infty[}(x)$ is smaller than ε. Since ε is arbitrary, q belongs to $\overline{\Psi_{[t,\infty[}(x)}$. □

Remark 3.10. *For a general differential inclusion with an upper semi-continuous bounded right-hand side with compact convex values, the omega limit set of a point is not (in general) strongly positively invariant, see e.g. [3].*

Proposition 3.11 (Properties of the accessible set)**.** *The set Γ satisfies the following:*

(i) $\Gamma = \bigcap_{x \in M} \omega_\Psi(x)$,

(ii) $\Gamma = \omega_\Psi(p)$ *for all* $p \in \Gamma$,

(iii) Γ *is compact, connected, invariant and strongly positively invariant under* Ψ,

(iv) *either* Γ *has empty interior or its interior is dense in* Γ.

Proof. **(i)** : Let $x \in M$ and $y \in \Psi_t(x)$. Then $\gamma^+(y) \subset \Psi_{[t,\infty]}(x)$. Hence $\Gamma \subset \overline{\Psi_{[t,\infty[}(x)}$ for all x. This proves that $\Gamma \subset \bigcap_{x\in M} \omega_\Psi(x)$. Conversely let $p \in \bigcap_{x\in M} \omega_\Psi(x)$. Then, for all $t > 0$ and $x \in M$, $p \in \overline{\Psi_{[t,\infty[}(x)} \subset \overline{\gamma^+(x)}$ where the latter inclusion follows from Lemma 3.5. This proves the converse inclusion.

(ii) : By Lemma 3.5, $\omega_\Psi(p) \subset \Gamma$. The converse inequality follows from **(i)**.

(iii) : This follows from **(ii)** and Lemma 3.9.

(iv) : Suppose $\mathrm{int}(\Gamma) \neq \emptyset$. Then there exists an open set $\mathcal{U} \subset \Gamma$ and $\bigcup_{\mathbf{t},\mathbf{i}} \mathbf{\Phi^i_t}(\mathcal{U})$ is an open subset of Γ dense in Γ. □

An equilibrium p for the flow Φ^1 is a point in M such that $F^1(p) = 0$. It is called an *attracting equilibrium* if there exists a neighborhood $\mathcal{U}$ of p such that

$$\lim_{t\to\infty} \|\Phi^1_t(x) - p\| = 0$$

uniformly in $x \in \mathcal{U}$. In this case, the *basin of attraction* of p is the open set

$$\mathcal{B}(p) = \left\{ x \in \mathbb{R}^d : \lim_{t\to\infty} \|\Phi^1_t(x) - p\| = 0 \right\}.$$

Proposition 3.12 (Case of an attracting equilibrium)**.** *Suppose the flow* Φ^1 *has an attracting equilibrium* p *with basin of attraction* $\mathcal{B}(p)$ *that intersects all orbits: for all* $x \in M \setminus \mathcal{B}(p)$, $\gamma^+(x) \cap \mathcal{B}(p) \neq \emptyset$. *Then*

(i) $\Gamma = \overline{\gamma^+(p)}$,

(ii) *If furthermore* $\Gamma \subset \mathcal{B}(p)$, *then* Γ *is contractible. In particular, it is simply connected.*

Proof. The proof of **(i)** is left to the reader. To prove **(ii)**, let $h : [0,1] \times \Gamma \to \Gamma$ be defined by

$$h(t,x) = \begin{cases} \Phi^1_{-\log(1-t)}(x) & \text{if } t < 1, \\ p & \text{if } t = 1. \end{cases}$$

It is easily seen that h is continuous. Hence the result. □

3.2.2 The accessible set and recurrence properties

In this section, we link the accessibility (which is a deterministic notion) to some recurrence properties for the embedded chain $\tilde{Z}$ and the continuous time process Z.

Proposition 3.13 (Returns near Γ — discrete case)**.** *Assume that* $\Gamma \neq \emptyset$. *Let* $p \in \Gamma$ *and* $\mathcal{U}$ *be a neighborhood of* p. *There exist* $m \in \mathbb{N}$ *and* $\delta > 0$ *such that for all* $i, j \in E$ *and* $x \in M$

$$\mathbb{P}_{x,i}\left[\tilde{Z}_m \in \mathcal{U} \times \{j\}\right] \geqslant \delta.$$

In particular,

$$\mathbb{P}_{x,i}\left[\exists l \in \mathbb{N},\ \tilde{Z}_l \in \mathcal{U} \times \{j\}\right] = 1.$$

Note that, in the previous proposition, the same discrete time m works for all $x \in M$. In the continuous time framework, one common time t does not suffice in general; however one can prove a similar statement if one allows a finite number of times:

Proposition 3.14 (Returns near Γ — continuous case)**.** *Assume that $\Gamma \neq \emptyset$. Let $p \in \Gamma$ and $\mathcal{U}$ be a neighborhood of p. There exist $N \in \mathbb{N}$, a finite open covering $\mathcal{O}^1, \dots \mathcal{O}^N$ of M, N times $t_1, \dots, t_N > 0$ and $\delta > 0$ such that, for all $i, j \in E$ and $x \in \mathcal{O}^k$,*

$$\mathbb{P}_{x,i}\left[Z_{t_k} \in \mathcal{U} \times \{j\}\right] \geqslant \delta.$$

Since Γ is positively invariant under each flow Φ^i we deduce from Proposition 3.14 the following result.

Corollary 3.15. *Assume Γ has non empty interior. Then*

$$\mathbb{P}_{x,i}\left[\exists t_0 \geqslant 0, \forall t \geqslant t_0, Z_t \in \Gamma \times E\right] = 1.$$

Propositions 3.13 and 3.14 are direct consequences of Lemma 3.2 and of the following technical result.

Lemma 3.16. *Assume $\Gamma \neq \emptyset$. Let $p \in \Gamma$, $\mathcal{U}$ be a neighborhood of p and $i, j \in E$. There exist $m \in \mathbb{N}^*, \varepsilon, \beta > 0$, finite sequences $(\mathbf{i}^1, \mathbf{u}^1) \dots, (\mathbf{i}^N, \mathbf{u}^N) \in \mathbb{T}_m^{ij}$ and an open covering $\mathcal{O}^1, \dots, \mathcal{O}^N$ of M (i.e. $M = \mathcal{O}^1 \cup \dots \cup \mathcal{O}^N$) such that for all $x \in M$ and $\tau \in \mathbb{R}_+^m$:*

$$x \in \mathcal{O}^k \text{ and } \left\|\tau - \mathbf{u}^k\right\| \leqslant \varepsilon \quad \Longrightarrow \quad \mathbf{\Phi}_\tau^{\mathbf{i}^k}(x) \in \mathcal{U} \text{ and } p(x, \mathbf{i}^k, \tau) \geqslant \beta.$$

Furthermore, m, ε and β can be chosen independent of $i, j \in E$.

Proof. Fix i and j, and let $\mathcal{V}$ be a neighborhood of p with closure $\overline{\mathcal{V}} \subset \mathcal{U}$. For all $\beta > 0$, define the open sets

$$\mathcal{O}(\mathbf{i}, \mathbf{u}, \beta) = \{x \in M : \mathbf{\Phi}_{\mathbf{u}}^{\mathbf{i}}(x) \in \mathcal{V}, p(x, \mathbf{i}, \mathbf{u}) > \beta\}.$$

By (15), one has

$$M = \bigcup_n \left(\bigcup_{\beta > 0} \bigcup_{(\mathbf{i},\mathbf{u}) \in \mathbb{T}_n^{ij}} \mathcal{O}(\mathbf{i}, \mathbf{u}, \beta) \right) \tag{16}$$

Now, since one can add "false" jumps (i does not change and u equals 0) to $(\mathbf{i}, \mathbf{u})$ without changing $\mathbf{\Phi}_{\mathbf{u}}^{\mathbf{i}}(x)$,

$$\forall (\mathbf{i}, \mathbf{u}) \in \mathbb{T}_n^{ij}, \forall n' \geqslant n, \forall \beta,\ \exists \beta' > 0, (\mathbf{i}', \mathbf{u}') \in \mathbb{T}_{n'}^{ij} \quad \text{such that} \quad \mathcal{O}(\mathbf{i}, \mathbf{u}, \beta) \subset \mathcal{O}(\mathbf{i}', \mathbf{u}', \beta')$$

(just add $n' - n$ false jumps at the beginning and let $\beta' = \beta(\inf_M Q(x, i, i))^{n'-n}$). Therefore the union over n in (16) is increasing: by compactness, there exists m such that

$$M \subset \bigcup_{\beta > 0} \bigcup_{(\mathbf{i},\mathbf{u}) \in \mathbb{T}_m^{ij}} \mathcal{O}(\mathbf{i}, \mathbf{u}, \beta).$$

Note that by monotonicity, m can be chosen uniformly over i and j. The union in β increases as β decreases, so by compactness again there exists β_0 (independent of i, j) such that

$$M \subset \bigcup_{(\mathbf{i},\mathbf{u}) \in \mathbb{T}_m^{ij}} \mathcal{O}(\mathbf{i}, \mathbf{u}, \beta_0).$$

A third invocation of compactness shows that for some finite N,

$$M \subset \bigcup_{k=1}^{N} \mathcal{O}_k,$$

where $\mathcal{O}^k = \mathcal{O}(\mathbf{i}^k, \mathbf{u}^k, \beta_0)$ for some $(\mathbf{i}^k, \mathbf{u}^k)$ in $\mathbb{T}_m^{ij}$. Since $\overline{\mathcal{V}} \subset \mathcal{U}$, the distance between $\mathcal{V}$ and $\mathcal{U}^c$ is positive (once more by compactness). Choosing ε small enough therefore guarantees that for $x \in \mathcal{O}^k$ and $\left\|\mathbf{v} - \mathbf{u}^k\right\| < \varepsilon$, $\mathbf{\Phi}_{\mathbf{i}}^{\mathbf{v}}(x) \in \mathcal{U}$. This concludes the proof. □

3.3 Support of invariant probabilities

The following proposition relates Γ to the support of the invariant measures of $\tilde{Z}$ or Z. We state and prove the result for $\tilde{Z}$ and rely on Lemma 2.6 for Z.

Proposition 3.17 (Accessible set and invariant measures)**.**

(i) *If $\Gamma \neq \emptyset$ then $\Gamma \times E \subset \mathrm{supp}(\mu)$ for all $\mu \in \tilde{\mathcal{P}}_{inv}$ and there exists $\mu \in \tilde{\mathcal{P}}_{inv}$ such that $\mathrm{supp}(\mu) = \Gamma \times E$.*

(ii) *If Γ has non empty interior, then $\Gamma \times E = \mathrm{supp}(\mu)$ for all $\mu \in \tilde{\mathcal{P}}_{inv}$.*

(iii) *Suppose that $\tilde{Z}$ is stable with invariant probability π, then $\mathrm{supp}(\pi) = \Gamma \times E$.*

Proof. **(i)** follows from Proposition 3.13. Also, since Γ is strongly positively invariant, there are invariant measures supported by $\Gamma \times E$. **(ii)** follows from **(i)** and Corollary 3.15. To prove **(iii)**, let $(p, i) \in \mathrm{supp}(\pi)$. Let $\mathcal{U}, \mathcal{V}$ be open neighborhoods of p with $\overline{\mathcal{U}} \subset \mathcal{V}$ compact. Let $0 \leqslant f \leqslant 1$ be a continuous function on M which is 1 on $\mathcal{U}$ and 0 outside $\mathcal{V}$ and let $\tilde{f}(x, j) = f(x)\delta_{j,i}$. Suppose $Z_0 = (x, j)$. Then, with probability one,

$$\liminf_{n\to\infty} \frac{1}{n} \sharp\Big\{1 \leqslant k \leqslant n \,:\, \tilde{Z}_k \in \mathcal{V} \times \{i\}\Big\} \geqslant \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \tilde{f}(\tilde{Z}_k) \geqslant \pi(\mathcal{U} \times \{i\}) > 0.$$

Hence $(\tilde{Z}_n)$ visits infinitely often $\mathcal{U} \times \{i\}$. In particular, $p \in \gamma^+(x)$. This proves that $\mathrm{supp}(\pi) \subset \Gamma \times E$. The converse statement follows from **(i)**. □

Remark 3.18. *The example given in Section 5.4 shows that the inclusion $\Gamma \times E \subset \mathrm{supp}(\mu)$ may be strict when Γ has empty interior. On the other hand, the condition that Γ has non empty interior is not sufficient to ensure uniqueness of the invariant probability since there exist smooth minimal flows that are not uniquely ergodic. An example of such a flow can be constructed on a 3-manifold by taking the suspension of an analytic minimal non uniquely ergodic diffeomorphism of the torus constructed by Furstenberg in [16] (see also [24]). As shown in [2] (see also Section 4) a sufficient condition to ensure uniqueness of the invariant probability is that the vector fields verify a Hörmander bracket property at some point in Γ.*

4 Absolute continuity and ergodicity

4.1 Absolute continuity of the law of the processes

Let x_0 be a point in M. The image of $u \mapsto \Phi_u(x_0)$ is a curve; one might expect that, if $i \neq j$, the image of $(s, t) \mapsto \Phi_t^j(\Phi_s^i(x_0))$ should be a surface. Going on composing the flow functions in this way, one might fill some neighbourhood of a point in $\mathbb{R}^d$.

Recall that, if $\mathbf{i}$ is a sequence of indices $\mathbf{i} = (i_0, \dots i_m)$ and $\mathbf{u}$ is a sequence of times $\mathbf{u} = (u_1, \dots u_m)$, $\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}} : M \to M$ is the composite map defined by (14).

Suppose that for some x_0 the map $\mathbf{v} \mapsto \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x_0)$ is a submersion at $\mathbf{u}$. Then the image of the Lebesgue measure on a neighbourhood of $\mathbf{u}$ is a measure on a neighbourhood of $\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x_0)$ equivalent to the Lebesgue measure. If the jump rates λ_i are constant functions, the probability $\mathbb{P}_{x_0,i_0}\left[\tilde{Z}_m \in \cdot \times \{i_m\}\right]$ is just the image by $\mathbf{u} \in \mathbb{R}^m \mapsto \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x_0) \in \mathbb{R}^d$ of the product of exponential laws on $\mathbb{R}^m$. We get that there exist $\mathcal{U}_0$ a neighborhood of x_0, $\mathcal{V}_0$ a neighborhood of $\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x_0)$, and a constant $c > 0$ such that:

$$\forall x \in \mathcal{U}_0, \quad \mathbb{P}_{x,i_0}\left[\tilde{Z}_m \in \cdot \times \{i_m\}\right] \geqslant c\lambda_{\mathbb{R}^d}(\cdot \cap \mathcal{V}_0).$$

Let us fix a $t > 0$ and consider now the function $\mathbf{v} \mapsto \Phi^{i_m}_{t-v_1-\dots-v_k}(\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x_0))$ defined on $v_1 + \dots + v_m < t$. For the same reason, if this function is a submersion at $\mathbf{u}$, then the law of X_t has an absolutely continuous part with respect to $\lambda_{\mathbb{R}^d}$.

The two following results state a stronger result (with a local uniformity with respect to initial and final positions) both for the embedded chain and the continuous time process.

Theorem 4.1 (Absolute continuity — discrete case)**.** *Let x_0 and y be two points in M and a sequence $(\mathbf{i}, \mathbf{u})$ such that $\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x_0) = y$. If $\mathbf{v} \mapsto \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x_0)$ is a submersion at $\mathbf{u}$, then there exist $\mathcal{U}_0$ a neighborhood of x_0, $\mathcal{V}$ a neighborhood of y, an integer m and a constant $c > 0$ such that*

$$\forall x \in \mathcal{U}_0, \ \forall i, j \in E, \quad \mathbb{P}_{x,i}\left[\tilde{Z}_m \in \cdot \times \{j\}\right] \geqslant c\lambda_{\mathbb{R}^d}(\cdot \cap \mathcal{V}). \tag{17}$$

Theorem 4.2 (Absolute continuity — continuous case)**.** *Let x_0 and y be two points in M and a sequence $(\mathbf{i}, \mathbf{u}) \in \mathbb{T}_m$ such that $\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x_0) = y$. If $\mathbf{v} \mapsto \mathbf{\Phi}^{i_m}_{s-(v_1+\dots+v_m)}\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x_0)$ is a submersion at $\mathbf{u}$ for some $s > u_1 + \dots + u_m$, then for all $t_0 > u_1 + \dots + u_m$, there exist $\mathcal{U}_0$ a neighborhood of x_0, $\mathcal{V}$ a neighborhood of y and two constants $c, \varepsilon > 0$ such that*

$$\forall x \in \mathcal{U}_0, \ \forall i, j \in E, \ \forall t \in [t_0, t_0 + \varepsilon], \quad \mathbb{P}_{x,i}\left[Z_t \in \cdot \times \{j\}\right] \geqslant c\lambda_{\mathbb{R}^d}(\cdot \cap \mathcal{V}). \tag{18}$$

The proofs of these results are postponed to Section 6.

Unfortunately, the hypotheses of these two theorems are not easy to check, since one needs to "solve" the flows. However, they translate to two very nice local conditions. To write down these conditions, we need a bit of additional notation. Let $\mathcal{F}_0$ the collection of vector fields $\{F^i : i \in E\}$. Let $\mathcal{F}_k = \mathcal{F}_{k-1} \cup \{[F^i, V], V \in \mathcal{F}_{k-1}\}$ (where $[F, G]$ stands for the Lie bracket of two vector fields F and G) and $\mathcal{F}_k(x)$ the vector space (included in $\mathbb{R}^d$) spanned by $\{V(x), V \in \mathcal{F}_k\}$.

Similarly, starting from $\mathcal{G}_0 = \{F^i - F^j, i \neq j\}$, we define $\mathcal{G}_k$ by taking Lie brackets with the vector fields $\{F^i : i \in E\}$, and $\mathcal{G}_k(x)$ the corresponding subspace of $\mathbb{R}^d$.

Definition 4.3. *We say that the* weak bracket condition *(resp.* strong bracket condition*) is satisfied at $x \in M$ if there exists k such that $\mathcal{F}_k(x) = \mathbb{R}^d$ (resp. $\mathcal{G}_k(x) = \mathbb{R}^d$).*

These two conditions are called A (for the stronger) and B (for the weaker) in [2]. Since $\mathcal{G}_k(x)$ is a subspace of $\mathcal{F}_k(x)$, the strong condition implies the weak one. The converse is false, a counter-example is given below in Section 5.1.

Theorem 4.4. *If the weak (resp. strong) bracket condition holds at $x_0 \in M$, then the conclusion of Theorem 4.1 (resp. Theorem 4.2) holds.*

This theorem is a version of Theorem 2 from [2] with an additional uniformity with respect to the initial point and the time t. Thanks to Theorems 4 and 5 in [2], one can deduce the hypotheses of Theorems 4.1 and 4.2 from the bracket conditions. The proofs in [2] are elegant but non-constructive; we give a more explicit proof in Section 7.

4.2 Ergodicity

4.2.1 The embedded chain

Theorem 4.5 (Convergence in total variation — discrete case)**.** *Suppose there exists $p \in \Gamma$ at which the weak bracket condition holds. Then the chain $\tilde{Z}$ admits a unique invariant probability $\tilde{\pi}$, absolutely continuous with respect to the Lebesgue measure $\lambda_{M\times E}$ on $M \times E$. Moreover, there exist two constants $c > 1$ and $\rho \in (0, 1)$ such that, for any $n \in \mathbb{N}$,*

$$\left\| \mathbb{P}\left[\tilde{Z}_n \in \cdot\right] - \tilde{\pi} \right\|_{TV} \leqslant c\rho^n$$

where $\|\cdot\|_{TV}$ stands for the total variation norm.

Proof. By Proposition 3.13 and Theorem 4.4, there exist a neighborhood $\mathcal{U}_0$ of p, integers m and K, a measure ψ on $M \times E$ absolutely continuous with respect to $\lambda_{M\times E}$ and $c > 0$ such that, setting $A = \mathcal{U}_0 \times E$,

(i) $\mathbb{P}_{x,i}\left[\tilde{Z}_m \in A\right] \geqslant \delta$ for all $(x, i) \in M \times E$,

(ii) $\mathbb{P}_{x,i}\left[\tilde{Z}_K \in \cdot\right] \geqslant c\psi(\cdot)$ for all $(x, i) \in A$.

These two properties make $\tilde{Z}$ an Harris chain with recurrence set A (see e.g. [15, Section 7.4]). It is recurrent, aperiodic (by **(i)**) and the first hitting time of A has geometric tail (by **(i)** again). Therefore, by usual arguments, two copies of $\tilde{Z}$ may be coupled in a time T that has geometric tail; this implies the exponential convergence in total variation toward its (necessarily unique) invariant probability $\tilde{\pi}$ (see e.g. the proof of Theorem 4.10 in [15, Section 7.4] or [23, Section I.3], for details).

Finally, observe that, by **(ii)** and Theorem 4.4, $\tilde{\pi} \geqslant \delta c\psi$. Therefore $\tilde{\pi}_{ac}$ (the absolutely continuous part of $\tilde{\pi}$ with respect to $\lambda_{M\times E}$) is non zero. Then, Proposition 2.8 ensures that $\tilde{\pi}$ is absolutely continuous with respect to $\lambda_{M\times E}$. □

As pointed out in [2, Theorem 1] (see also Theorem 2.4), under the hypotheses of Theorem 4.5, $(Z_t)_{t\geqslant 0}$ admits a unique invariant probability measure π, and π is absolutely continuous with respect to $\lambda_{M\times E}$. Moreover, under the hypotheses of Theorem 4.5, with probability one

$$\lim_{n\to\infty} \tilde{\Pi}_n = \tilde{\pi} \quad \text{and} \quad \lim_{t\to\infty} \Pi_t = \tilde{\pi}\tilde{K}.$$

Under the strong bracket assumption, we prove in the next section that the distribution of Z_t itself converges, and not only its empirical measure.

4.2.2 The continuous time process

Theorem 4.6 (Convergence in total variation — continuous case)**.** *Suppose that there is a point $p \in \Gamma$ at which the strong bracket condition is satisfied. Let π be the unique invariant probability measure of Z. Then there exist two constants $c > 1$ and $\alpha > 0$ such that, for any $t \geqslant 0$,*

$$\|\mathbb{P}\left[Z_t \in \cdot\right] - \pi\|_{TV} \leqslant ce^{-\alpha t}. \tag{19}$$

Proof. The proof consists in showing that there exist a neighborhood $\mathcal{U}$ of p, $t > 0$, $j \in E$ and $\beta > 0$ such that for all $x \in M$ and $i \in E$,

$$\mathbb{P}_{x,i}\left[Z_t \in \mathcal{U} \times \{j\}\right] \geqslant \beta. \tag{20}$$

This property and (18) ensure that two processes starting from anywhere can be coupled in some time t with positive probability. This, combined with Theorem 4.4, implies (19) by the usual coupling argument (see [23]).

Theorem 4.4 gives us two open sets $\mathcal{U}_0$, $\mathcal{V}$ (with $p \in \mathcal{U}_0$), a time s_0 and $\varepsilon > 0$ such that

$$\forall x \in \mathcal{U}_0, \forall t \in [s_0, s_0 + \varepsilon], \forall i, \forall j, \quad \mathbb{P}_{x,i}\left[Z_t \in \cdot \times \{j\}\right] \geqslant c\lambda_{\mathbb{R}^d}(\cdot \cap \mathcal{V}). \tag{21}$$

Moreover, we have seen in Proposition (3.14) that there exist m times $t_1, \ldots, t_m > 0$, m open sets $\mathcal{O}^1, \ldots, \mathcal{O}^m$ (covering M) and $\delta > 0$ such that, for all $i, j \in E$ and $x \in \mathcal{O}^k$,

$$\mathbb{P}_{x,i}\left[Z_{t_k} \in \mathcal{U}_0 \times \{j\}\right] \geqslant \delta. \tag{22}$$

We can suppose that $\mathcal{V}$ is included in one of the $\mathcal{O}^k$ (just shrink $\mathcal{V}$ if necessary). Then, there exists $s \in \{t_1, \ldots, t_m\}$ such that

$$\forall y \in \mathcal{V}, \forall i, \forall j, \quad \mathbb{P}_{y,i}\left[Z_s \in \mathcal{U}_0 \times \{j\}\right] \geqslant \beta_1. \tag{23}$$

From (21) and (23), we get a time $t_0 (= s_0 + s)$ and $\beta_2 > 0$, such that

$$\forall x \in \mathcal{U}_0, \forall t \in [t_0, t_0 + \varepsilon], \forall i, \forall j, \quad \mathbb{P}_{x,i}\left[Z_t \in \mathcal{U}_0 \times \{j\}\right] \geqslant \beta_2.$$

The Markov property then gives that, for every integer n, one has

$$\forall x \in \mathcal{U}_0, \forall t \in [nt_0, nt_0 + n\varepsilon], \forall i, \forall j, \quad \mathbb{P}_{x,i}\left[Z_t \in \mathcal{U}_0 \times \{j\}\right] \geqslant \beta_2^n. \tag{24}$$

We can suppose that t_1 is the largest of the m times $(t_k)_{1 \leqslant k \leqslant m}$. Let n be such that for every k between 1 and m there exists $v_k \in [0, n\varepsilon]$ such that $t_1 = t_k + v_k$. Take such numbers $(v_k)_{1 \leqslant k \leqslant m}$. If x belongs to $\mathcal{O}^k$, by (22) and (24), for all i and j,

$$\mathbb{P}_{x,i}\left[Z_{t_1 + nt_0} \in \mathcal{U}_0 \times \{j\}\right] = \mathbb{P}_{x,i}\left[Z_{t_k + nt_0 + v_k} \in \mathcal{U}_0 \times \{j\}\right] \geqslant \delta\beta_2^n.$$

This concludes the proof. □

5 Examples

5.1 On the torus

Consider the system defined on the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ by the constant vector fields $F^i = e^i$, where $(e_1, \ldots e_d)$ is the standard basis on $\mathbb{R}^d$. Then, as argued in [2], the weak bracket condition holds everywhere, and the strong condition does not hold. Therefore the chain $\tilde{Z}$ is ergodic and converges exponentially fast, the empirical means of $(\tilde{Z}_n)$ and (Z_t) converge, but the law of Z_t is singular with respect to the invariant measure for any $t > 0$ provided it is true for $t = 0$.

5.2 Two planar linear flows

Let A be a 2×2 real matrix whose eigenvalues η_1, η_2 have negative real parts. Set $E = \{0, 1\}$ and consider the process defined on $\mathbb{R}^2 \times E$ by

$$F^0(x) = Ax \quad \text{and} \quad F^1(x) = A(x - a)$$

for some $a \in \mathbb{R}^2$. The associated flows are $\Phi_t^0(x) = e^{tA}x$ and $\Phi_t^1(x) = e^{tA}(x - a) + a$. Each flow admits a unique equilibrium (which is attracting): 0 and a respectively.

First note that, by using the Jordan decomposition of A, it is possible to find a scalar product $\langle \cdot \rangle$ on $\mathbb{R}^2$ (depending on A) and some number $0 < \alpha \leqslant \min(-\mathrm{Re}(\eta_1), -\mathrm{Re}(\eta_2))$ such that $\langle Ax, x\rangle \leqslant -\alpha\langle x, x\rangle$. Therefore

$$\langle A(x-a), x\rangle \leqslant -\alpha\langle x, x\rangle - \langle Aa, x\rangle \leqslant \|x\|(-\alpha\|x\| + \|Aa\|).$$

This shows that, for $R > \|Aa\|/\alpha$, the ball $M = \{x \in \mathbb{R}^2, \|x\| \leqslant R\}$ is positively invariant by Φ^0 and Φ^1. Moreover every solution to the differential inclusion induced by $\{F^0, F^1\}$ eventually enters M. In particular $M \times E$ is an absorbing set for the process Z.

Another remark that will be useful in our analysis is that

$$\det(F^0(x), F^1(x)) = \det(A)\det(a, x),$$

so that

$$\det(F^0(x), F^1(x)) > 0 \ \ (\text{resp.} = 0) \Leftrightarrow \det(a, x) > 0 \ \ (\text{resp.} = 0). \tag{25}$$

Case 1: a is an eigenvector

If a is an eigenvector of A, then the line $\mathbb{R}a$ is invariant by both flows, so that

$$\Gamma = \overline{\gamma^+(0)} = [0, a]$$

and there is a unique invariant probability π (and its support is Γ by Proposition 3.17). Indeed, it is easily seen that Γ is an attractor for the set-valued dynamics induced by F^0 and F^1. Therefore the support of every invariant measure equals Γ. If we consider the process restricted to Γ, it becomes one-dimensional and the strong bracket condition holds, proving uniqueness.

Remark 5.1. *If $X_0 \notin \mathbb{R}a$, X will never reach Γ. As a consequence, the law of X_t and π are singular for any $t \geqslant 0$. In particular, their total variation distance is constant, equal to 1. Note also that, the strong bracket condition being satisfied everywhere except on $\mathbb{R}a$, the law of X_t at any positive finite time has a non trivial absolutely continuous part.*

Remark 5.2. *Consider the following example: $A = -I$, $a = (1, 0)$ and $\mathbb{R}a$ is identified to $\mathbb{R}$. If the jump rates are constant and equal to λ, it is easy to check (see [20, 27]) that the invariant measure μ on $[0, 1] \times \{0, 1\}$ is given by:*

$$\mu = \frac{1}{2}\left(\mu_0 \times \delta_0 + \mu_1 \times \delta_1\right),$$

where μ_0 and μ_1 are Beta laws on $[0, 1]$:

$$\mu_0(dx) = C_\lambda x^{\lambda-1}(1-x)^\lambda\, dx,$$
$$\mu_1(dx) = C_\lambda x^\lambda(1-x)^{\lambda-1}\, dx.$$

In particular, this example shows that the density of the invariant measure (with respect to the Lebesgue measure) may be unbounded: when the jump rate λ is smaller than 1, the densities blow up at 0 and 1.

Case 2: Eigenvalues are reals and a is not an eigenvector

Suppose that the two eigenvalues η_1 and η_2 of A are negative real numbers and that a is not an eigenvector of A.

Let $\gamma_0 = \{\Phi_t^0(a),\, t \geqslant 0\}$ and $\gamma_1 = \{\Phi_t^1(0),\, t \geqslant 0\}$. Note that γ_1 and γ_0 are image of each other by the transformation $T(x) = a - x$. The curve γ_0 (respectively γ_1) crosses the line $\mathbb{R}a$ only at point a (respectively 0). Otherwise, the trajectory $t \mapsto \Phi_t^0(a)$ would have to cross the line $\mathrm{Ker}(A - \lambda_1 I)$ which is invariant. This makes the curve $\gamma = \gamma_0 \cup \gamma_1$ a simple closed curve in $\mathbb{R}^2$ crossing $\mathbb{R}a$ at 0 and a. By Jordan curve Theorem, $\mathbb{R}^2 \setminus \gamma = \mathcal{B} \cup \mathcal{U}$ where $\mathcal{B}$ is a bounded component and $\mathcal{U}$ an unbounded one. We claim that

$$\Gamma = \overline{\mathcal{B}}.$$

To prove this claim, observe that thanks to (25), F^0 and F^1 both point inward $\mathcal{B}$ at every point of γ. This makes $\overline{\mathcal{B}}$ positively invariant by Φ^0 and Φ^1. Thus $\Gamma \subset \overline{\mathcal{B}}$. Conversely, $\gamma \subset \Gamma$ (because 0 and a are accessible from everywhere). If $x \in \mathcal{B}$ there exists $s > 0$ such that $\Phi_{-s}^0(x) \in \gamma$ (because $\lim_{t\to-\infty} |\Phi_t^0(x)| = +\infty$) and necessarily $\Phi_{-s}^0(x) \in \gamma_1$. This proves that $x \in \gamma^+(0)$. Finally note that the strong bracket condition is verified in $\Gamma \setminus \mathbb{R}a$, proving uniqueness and absolute continuity of the invariant probability.

Remark 5.3. *Note that if the jump rates are small, the situation is similar to the one described in Remark 5.2: the process spends a large amount of time near the attractive points, and the density is unbounded at these points. By the way, this is also the case on the boundary of* Γ.

Case 3: Eigenvalues are complex conjugate

Suppose now that the eigenvalues have a nonzero imaginary part. By Jordan decomposition, it is easily seen that trajectories of Φ^i converge in spiralling, so that the mappings $\tau^i(x) = \inf\{t > 0 :\ \Phi_t^i(x) \in \mathbb{R}a\}$ and $h^i(x) = \Phi^i_{\tau^i(x)}$ are well defined and continuous. Let $H : \mathbb{R}a \to \mathbb{R}a$ be the map $h^0 \circ h^1$ restricted to $\mathbb{R}a$. Since two different trajectories of the same flow have empty intersection, the sequence $x_n = H^n(0)$ is decreasing (for the ordering on $\mathbb{R}a$ inherited from $\mathbb{R}$). Being bounded (recall that M is compact and positively invariant), it converges to $x^* \in \mathbb{R}a$ such that $x^* = H(x^*)$. Let now $\gamma^0 = \{\Phi_t^1(x^*),\ 0 \leqslant t \leqslant \tau^1(x^*)\}, \gamma^1 = \{\Phi_t^0(h^1(x^*)),\ 0 \leqslant t \leqslant \tau^0(h^1(x^*))\}$ and $\gamma = \gamma^0 \cup \gamma^1$. Reasoning as previously shows that Γ is the bounded component of $\mathbb{R}^2 \setminus \gamma$ and that there is a unique invariant and absolutely continuous invariant probability.

We illustrate this situation in Figure 1, with

$$A = \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{26}$$

Remark 5.4. *Once again, if the jump rates are small, then the density is unbounded at 0 but also on the set*

$$\Big\{\Phi_t^1(0),\ t \geqslant 0\Big\} \cup \Big\{\Phi_t^0(a),\ t \geqslant 0\Big\}.$$

5.3 A simple criterion for the accessible set to have a non empty interior

Here is a simple criteria in dimension 2 that ensures that Γ has a non empty interior.

Proposition 5.5. *Assume that $M \subset \mathbb{R}^2$, $E = \{0, 1\}$ and that F^1 has a globally attracting equilibrium p such that the eigenvalues of $DF^1(p)$ have negative real parts and that $F^0(p) \neq 0$. Then p lies in the interior of* Γ.

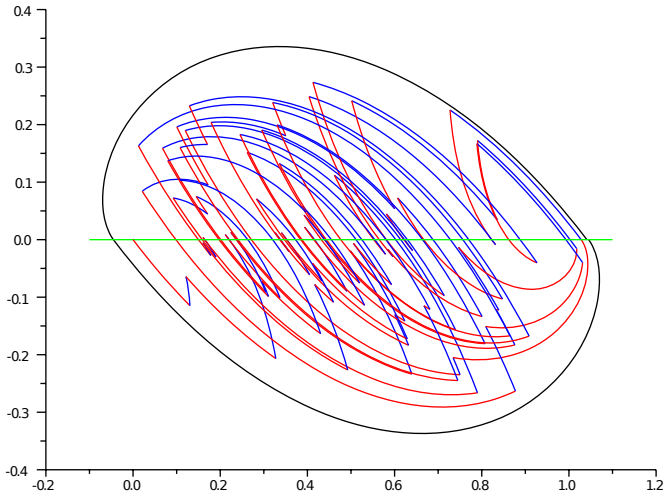


Figure 1: A sample path of X_t (red and blue lines) and the boundary of the support of the invariant probability (black line) for the third case in Section 5.2, where A and a are given by (26).

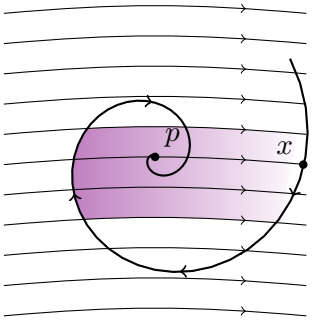


Figure 2: The flows near the attracting point p in Proposition 5.5.

Proof. Proposition 3.12 ensures that p belongs to Γ. As illustrated by Figure 2, from the equilibrium p, one can follow F^0 and reach x, then follow F^1, and switch back to F^0 to reach any point in the shaded region. □

5.4 Knowing the flows is not enough

In this section we study a PDMP on $\mathbb{R}^2 \times \{0,1\}$ such that the strong bracket condition holds everywhere except on Γ and which may have one or three ergodic invariant probability measures, depending on the jump rates of the discrete part of the process.

This model has been suggested by O. Radulescu. The continuous part of the process takes its values on $\mathbb{R}^2$ whereas its discrete part belongs to $\{0,1\}$. For sake of simplicity we will denote (in a different way than in the beginning of the paper) by $(X_t, Y_t) \in \mathbb{R}^2$ the continuous component. The discrete component $(I_t)_{t\geqslant 0}$ is a continuous time Markov chain on $E = \{0,1\}$ with jump rates $(\lambda_i)_{i\in E}$. Let $\alpha > 0$. The two vector fields F^0 and F^1 are

given by

$$F^0(x,y) = \begin{pmatrix} -x+\alpha \\ -y+\alpha \end{pmatrix} \quad \text{and} \quad F^1(x,y) = \begin{pmatrix} -x + \dfrac{\alpha}{1+y^2} \\ -y + \dfrac{\alpha}{1+x^2} \end{pmatrix}$$

with $(x,y) \in \mathbb{R}^2$. Notice that the quarter plane $(0,+\infty)^2$ is positively invariant by Φ^0 and Φ^1. See Figure 3. In the sequel we assume that (X_0, Y_0) belongs to $(0,+\infty)^2$.

5.4.1 Properties of the two vector fields

Obviously, the vector fields F^0 has a unique stable point (α,α). The description of F^1 is more involved and depends on α.

Lemma 5.6. *Let us define*

$$a = \frac{\alpha + \sqrt{|\alpha^2 - 4|}}{2} \quad \textit{and} \quad b = \left(\frac{\sqrt{4/27+\alpha^2}+\alpha}{2}\right)^{1/3} - \left(\frac{\sqrt{4/27+\alpha^2}-\alpha}{2}\right)^{1/3}. \quad (27)$$

Notice that b is positive and is the unique real solution of $b^3+b=\alpha$. One has

- *if $\alpha \leqslant 2$, then F^1 admits a unique critical point (b,b) and it is stable,*
- *if $\alpha > 2$, then F^1 admits three critical points: (b,b) is unstable whereas (a,a^{-1}) and (a^{-1},a) are stable.*

Proof. If (x,y) is a critical point of F^1 then (x,y) is solution of

$$\begin{cases} x(1+y^2) = \alpha \\ y(1+x^2) = \alpha. \end{cases}$$

As a consequence, x is solution of

$$0 = x^5 - \alpha x^4 + 2x^3 - 2\alpha x^2 + (1+\alpha^2)x - \alpha = (x^2 - \alpha x + 1)(x^3 + x - \alpha).$$

The equation $x^3+x-\alpha=0$ admits a unique real solution b given by (27). It belongs to $(0,\alpha)$. Obviously, if $\alpha \leqslant 2$, (b,b) is the unique critical point of F^1 whereas, if $\alpha > 2$ then a and a^{-1} are the roots of $x^2 - \alpha x + 1 = 0$ and F^1 admits the three critical points: (b,b), (a,a^{-1}) and (a^{-1},a). Let us have a look to the stability of (b,b). The eigenvalues of $\mathrm{Jac}(F^1)(b,b)$ are given by

$$\eta_1 = -3 + \frac{2b}{\alpha} = -1 - 2\frac{\alpha - b}{\alpha} \quad \text{and} \quad \eta_2 = 1 - \frac{2b}{\alpha} = \frac{b^3 - b}{\alpha}$$

and are respectively associated to the eigenvectors $(1,1)$ and $(1,-1)$. Since $b < \alpha$, η_1 is smaller than -1. Moreover, η_2 has the same sign as $b-1$ *i.e.* the same sign as $\alpha - 2$. As a conclusion, (b,b) is stable (resp. unstable) if $\alpha < 2$ (resp. $\alpha > 2$).

Assume now that $\alpha > 2$. Then $\mathrm{Jac}(F^1)(a,a^{-1})$ has two negative eigenvalues $-1 \pm 2\alpha^{-1}$. Then, the critical points (a,a^{-1}) and (a^{-1},a) are stable. □

In the sequel, we assume that $\alpha > 2$. One can easily check that the sets

$$D = \{(x,x) \,:\, x > 0\}, \quad L = \{(x,y) \,:\, 0 < y < x\}, \quad U = \{(x,y) \,:\, 0 < x < y\}$$

are strongly positively invariant by Φ^0 and Φ^1. Moreover, thanks to Proposition 3.12, the accessible set of (X,Y) is

$$\Gamma = \{(x,x) \,:\, x \in [b,\alpha]\}.$$

In the sequel, we prove that Γ may, or may not, be the set of all recurrent points, depending on the jump rates λ_0 and λ_1.

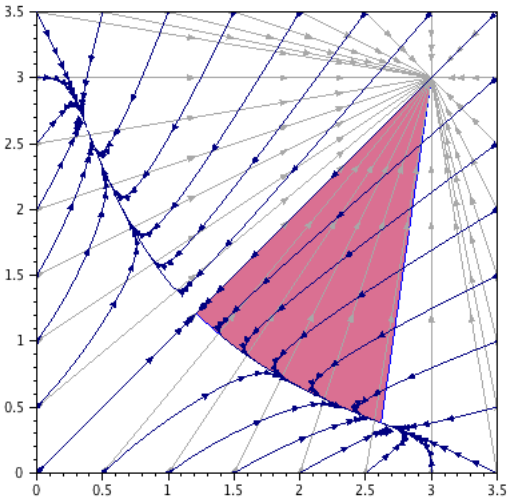


Figure 3: Vector fields F^0 (grey lines) and F^1 (blue lines) of the example in Section 5.4 with $\alpha = 3$. The shaded region is $\overline{\gamma^+(a, a^{-1})} \times E$.

Proposition 5.7. *If $\lambda_1 > \lambda_0(c\alpha - 1)$, with $c = 3\sqrt{3}/8$ then (X, Y, I) admits a unique invariant measure and its support is $\Gamma \times E$.*

If λ_1/λ_0 is small enough, then (X, Y, I) admits three ergodic measures and they are supported by

$$\Gamma \times E = \overline{\gamma^+(\alpha, \alpha)} \times E, \quad \overline{\gamma^+(a, a^{-1})} \times E \quad \text{and} \quad \overline{\gamma^+(a^{-1}, a)} \times E.$$

Remark 5.8. *This dichotomy is essentially due to the fact that the stable manifold $\{(x, x) : x \in \mathbb{R}\}$ of the unstable critical point (b, b) of F^1 is strongly positively invariant by Φ^0 (see Figure 3). Moreover, $\overline{\gamma^+(a, a^{-1})} \times E$ can be written as the union of the segment $[(a, a^{-1}), (\alpha, \alpha)]$ and Γ and the unstable manifold (included in L) of (b, b) for Φ^1.*

The following two sections are dedicated to the proof of Proposition 5.7.

5.4.2 Transience

The goal of this section is to prove the first part of Proposition 5.7.

Lemma 5.9. *Assume that $(X_0, Y_0) \in L$. Then, for any $t > 0$,*

$$0 < X_t - Y_t \leqslant (X_0 - Y_0) \exp\left(-\int_0^t \alpha(I_s)\, ds\right),$$

with $\alpha(0) = 1$ and $\alpha(1) = 1 - c\alpha < 0$ with $c = (3/8)\sqrt{3}$.

Proof. If $I_t = 0$ then

$$\frac{d}{dt}(X_t - Y_t) = -(X_t - Y_t).$$

On the other hand, if $I_t = 1$ then

$$\begin{aligned}\frac{d}{dt}(X_t - Y_t) &= -(X_t - Y_t) + \alpha \frac{X_t^2 - Y_t^2}{(1+X_t^2)(1+Y_t^2)} \\ &= -(1 - \alpha h(X_t, Y_t))(X_t - Y_t)\end{aligned}$$

where the function h is defined on $[0,\infty)^2$ by

$$h(x,y) = \frac{x+y}{(1+x^2)(1+y^2)}.$$

The unique critical point of h on $[0,\infty)^2$ is $(1/\sqrt{3}, 1/\sqrt{3})$ and h reaches its maximum at this point:

$$c := \sup_{x,y>0} h(x,y) = \frac{3\sqrt{3}}{8}.$$

As a consequence, for any $t \geqslant 0$,

$$\frac{d}{dt}(X_t - Y_t) \leqslant -\alpha(I_t)(X_t - Y_t) \quad \text{where} \quad \begin{cases}\alpha(0) = 1, \\ \alpha(1) = 1 - c\alpha.\end{cases}$$

Integrating this relation concludes the proof. □

Corollary 5.10. *Assume that $(X_0, Y_0) \in L$. If $\lambda_1 > \lambda_0(c\alpha - 1)$ then (X_t, Y_t) converges exponentially fast to D almost surely. More precisely,*

$$\limsup_{t\to\infty} \frac{1}{t} \log(X_t - Y_t) \leqslant -\frac{\lambda_1 - (c\alpha - 1)\lambda_0}{\lambda_0 + \lambda_1} < 0 \quad a.s. \tag{28}$$

In particular, the process (X, Y, I) admits a unique invariant measure μ. Its support is the set

$$S = \{(x, x, i) \,:\, x \in [b, \alpha],\ i \in \{0, 1\}\}.$$

Proof. The ergodic theorem for the Markov process $(I_t)_{t\geqslant 0}$ ensures that

$$\frac{1}{t}\int_0^t \alpha(I_s)\, ds \xrightarrow[t\to\infty]{a.s.} \int \alpha(i) d\nu(i)$$

where the invariant measure ν of the process $(I_t)_{t\geqslant 0}$ is the Bernoulli measure with parameter $\lambda_0/(\lambda_0 + \lambda_1)$. The upper bound (28) is a straightforward consequence of Lemma 5.9. This ensures that the sets L and U are transient. □

5.4.3 Recurrence

In this section, we aim at proving the second part of Proposition 5.7. Let us define the following new variables:

$$U_t = \frac{X_t + Y_t}{2} \quad \text{and} \quad V_t = \frac{X_t - Y_t}{2}.$$

Of course (U, V, I) is still a PDMP. If

$$\frac{d}{dt}\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = F^1(X_t, Y_t) \quad \text{then} \quad \frac{d}{dt}\begin{pmatrix} U_t \\ V_t \end{pmatrix} = G^1(U_t, V_t),$$

with

$$G^1(u,v) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} F^1(u+v, u-v) = \begin{pmatrix} -u + \dfrac{\alpha(1+u^2+v^2)}{(1+(u+v)^2)(1+(u-v)^2)} \\ -v + \dfrac{2\alpha uv}{(1+(u+v)^2)(1+(u-v)^2)} \end{pmatrix}.$$

Corollary 5.10 ensures that, if λ_1/λ_0 is large enough, then V_t goes to 0 exponentially fast. Let us show that this is no longer true if λ_1/λ_0 is small enough. Let $\varepsilon > 0$. Assume that, with positive probability, $V_t \in (0, \varepsilon)$ for any $t \geqslant 0$. Then, for any time $t \geqslant 0$, $(U_t, V_t) \in [b, \alpha] \times [0, \varepsilon]$. Indeed, one can show that $U_t \in [b, \alpha]$ for any $t \geqslant 0$ as soon as it is true at $t = 0$. The following lemma states that the vector fields G^1 can be compared to a vector fields H^1 (which is simpler to study).

Lemma 5.11. *Assume that $(u, v) \in [b, \alpha] \times [0, \varepsilon]$. Then there exist $u_c \in (b, \alpha)$ and $K, \delta, \gamma, \tilde{\gamma} > 0$ (that do not depend on ε) such that $b_\varepsilon = b + K\varepsilon^2$ and*

$$G_1^1(u,v) \leqslant H_1^1(u,v) \quad \textit{with} \quad H_1^1(u,v) = -\delta(u - b_\varepsilon).$$

and

$$G_2^1(u,v) \geqslant H_2^1(u,v) \quad \textit{with} \quad H_2^1(u,v) = \Big((\gamma + \tilde{\gamma})\mathbb{1}_{\{u \leqslant u_c\}} - \tilde{\gamma}\Big)v.$$

Proof. Notice firstly that there exists $c > 0$ such that

$$\forall (u,v) \in [b,\alpha] \times [0,\varepsilon], \quad \Big|(1+(u+v)^2)(1+(u-v)^2) - (1+u^2)^2\Big| \leqslant c\varepsilon^2. \tag{29}$$

Thus, using that $u^3 + u - \alpha = (u-b)(u^2 + bu + \alpha/b)$ and $u > b$ we get that

$$\begin{aligned} G_1^1(u,v) &\leqslant -u + \frac{\alpha}{1+u^2} + c\varepsilon^2 \\ &\leqslant -(u-b)\frac{u^2 + bu + \alpha/b}{1+u^2} + c\varepsilon^2 \\ &\leqslant -(u-b)\frac{2b^2 + \alpha/b}{1+\alpha^2} + c\varepsilon^2. \end{aligned}$$

We get the desired upper bound for G_1^1 with

$$\delta = \frac{2b^2 + \alpha/b}{1+\alpha^2}, \quad K = c/\delta \quad \text{and} \quad b_\varepsilon = b + K\varepsilon^2.$$

Similarly, Equation (29) ensures that

$$G_2^1(u,v) \geqslant v g(u) \quad \text{with} \quad g(u) = \frac{2\alpha u}{(1+u^2)^2} - 1 - c\varepsilon^2.$$

Obviously, if ε is small enough, $g(b) > 0$, $g(\alpha) < 0$ and g is decreasing on $[b, \alpha]$. Thus, if $\tilde{u}$ is the unique zero of g on (b, α), then one can choose

$$u_c = \frac{\tilde{u} + b}{2}, \quad \gamma = g(u_c) \quad \text{and} \quad \tilde{\gamma} = g(\alpha).$$

To get a simpler bound in the sequel we can even set $\tilde{\gamma} = g(\alpha) \vee 1$. □

Finally, define $H_1^0(u,v) = G_1^0(u,v) = -(u-\alpha)$ and $H_2^0(u,v) = G_2^0(u,v) = -v$ and introduce the PDMP $(\tilde{U}, \tilde{V}, \tilde{I})$ where $\tilde{I} = I$ is the switching process of (U, V, I) and $(\tilde{U}, \tilde{V})$ is driven by H^0 and H^1 instead of G^0 and G^1. From Lemma 5.11, we get that

$$\forall t \geqslant 0, \quad U_t \leqslant \tilde{U}_t \quad \text{and} \quad \tilde{V}_t \leqslant V_t$$

assuming that $(\tilde{U}_0, \tilde{V}_0, \tilde{I}_0) = (U_0, V_0, I_0)$. The last step is to study briefly the process $(\tilde{U}, \tilde{V}, \tilde{I})$. From the definition of the vector fields that drive $(\tilde{U}, \tilde{V}, \tilde{I})$, one has

$$\frac{d}{dt}\tilde{V}_t = \begin{cases} -\tilde{V}_t & \text{if } I_t = 0, \\ ((\gamma + \tilde{\gamma})\mathbb{1}_{\{\tilde{U}_t \leqslant u_c\}} - \tilde{\gamma})V_t & \text{if } I_t = 1. \end{cases}$$

This ensures that

$$\frac{1}{t}\log\frac{\tilde{V}_t}{\tilde{V}_0} \geqslant \frac{1}{t}\int_0^t ((\gamma + \tilde{\gamma})\mathbb{1}_{\{I_s = 1, \tilde{U}_s \leqslant u_c\}} - \tilde{\gamma})\, ds \tag{30}$$

since $\tilde{\gamma} \geqslant 1$. If λ_1/λ_0 is small enough, then $(I_s, \tilde{U}_s)$ spends an arbitrary large amount of time near $(1, b_\varepsilon)$ (and it can be assumed that $b_\varepsilon < u_c$ if ε is small enough). Thus, the right-hand side of (30) converges almost surely to a positive limit as soon as λ_1/λ_0 is small enough. As a consequence, V cannot be bounded by ε forever. The Markov property ensures that (X, Y) can reach any neighborhood of (a, a^{-1}) with probability 1 and thus, (a, a^{-1}) belongs to the support of the invariant measure. This concludes the proof of the second part of Proposition 5.7.

6 Absolute continuity — proofs of global criteria

This section is devoted to the proof of Theorems 4.1 and 4.2. The main idea of the proof has already been given before the statements; the main difficulty lies in providing estimates that are (locally) uniform in the starting point (x, i), the region of endpoints $\mathcal{V} \times E$, and the discrete time m or the continuous time t.

After seeing in Section 6.1 what the submersion hypothesis means in terms of vector fields, we establish in Section 6.2 a parametrized version of the local inversion lemma. This provides the uniformity in the continuous part x of the starting point, and enables us to prove in Section 6.3 a weaker version of Theorem 4.1. In Section 6.4 we show how to prove the result in its full strength. Finally, the fixed time result (Theorem 4.2) is proved in Section 6.5.

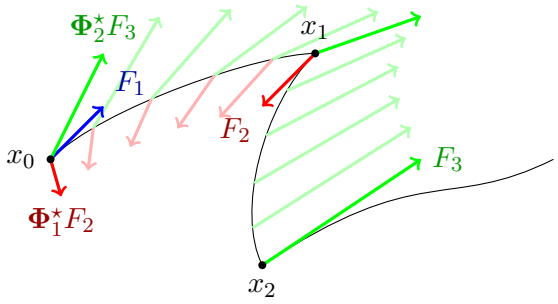
6.1 Submersions, vector fields and pullbacks

Before going into the details of the proof, let us see how one can interpret the submersion hypotheses of our regularity theorems.

Recall that, for x and $\mathbf{i}$ fixed, we are interested in the map

$$\begin{aligned} \psi : \mathbb{R}^m &\to \mathbb{R}^d \\ \mathbf{v} &\mapsto \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x). \end{aligned}$$

To see if this is a submersion at $\mathbf{u}$, we compute the partial derivatives with respect to the v_i: these are elements of $T_{x_m}M$, and ψ is a submersion if and only if these m vectors span $T_{x_m}M$. This is the case if and only if their inverse image by $D\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}$ span $T_{x_0}M$. An easy computation (see also Figure 4) shows that these vectors are given by:



In this picture $(i_0, i_1, i_2) = (1, 2, 3)$. The trajectory starts at x_0, and follows $F^{i_0} = F^1$ for a time u_1. At the first jump, it starts following $F^{i_1} = F^2$; we pull this tangent vector, depicted in red, back to x_0. The next (green) tangent vector $F^{i_2} = F^3$ (at x_2) has to be pulled back by the two flows. If the three tangent vectors we obtain at x_0 span $T_{x_0}M$, $\mathbf{v} \mapsto \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x_0)$ is a submersion.

Figure 4: The global condition

$$\tilde{\mathcal{C}}(\mathbf{i}, \mathbf{u}) = \left\{ F^{i_0}(x_0),\ \mathbf{\Phi}_1^\star F^{i_1}(x_0),\ \ldots,\ \mathbf{\Phi}_m^\star F^{i_m}(x_0) \right\}, \tag{31}$$

where $\mathbf{\Phi}_k^\star$ is the composite pullback:

$$\mathbf{\Phi}_k^\star = \Phi_{u_1}^{i_0,\star} \circ \cdots \circ \Phi_{u_k}^{i_{k-1},\star}. \tag{32}$$

Note that $\mathbf{\Phi}_k^\star$ depends on $\mathbf{i}$ and $\mathbf{u}$, but we hide this dependence for the sake of readability.

6.2 Parametrized local inversion

Let us first prove a "uniform" local inversion lemma, for functions of $\mathbf{t}$ that depend on a parameter x.

Remark 6.1. *Even if x lives in some $\mathbb{R}^d$, we do not write it in boldface, for the sake of coherence with the rest of the paper.*

Lemma 6.2. *Let d and m be two integers, and let f be a $\mathcal{C}^1$ map from $\mathbb{R}^m \times \mathbb{R}^d$ to $\mathbb{R}^m$,*

$$f : (\mathbf{t}, x) \mapsto f(\mathbf{t}, x) = f_x(\mathbf{t}).$$

For any fixed x, f_x maps $\mathbb{R}^m$ to itself; we denote its derivative at $\mathbf{t}$ by $(Df_x)_{\mathbf{t}}$. Suppose that, for some points x_0 and $\mathbf{t}_0$, $(Df_{x_0})_{\mathbf{t}_0}$ is invertible. Then we can find a neighborhood $J \subset \mathbb{R}^d$ of x_0, an open set $I \subset \mathbb{R}^m$ and, for all $x \in J$, an open set $W_x \subset \mathbb{R}^m$, such that:

$$\tilde{f}_x : \begin{cases} W_x & \to I, \\ \mathbf{t} & \mapsto f_x(\mathbf{t}) \end{cases}$$

is a diffeomorphism. Moreover, for any integer $k \leqslant m$, and any neighborhood W of $\mathbf{t}_0$, we can choose I, J and the W_x so that:

i) I is a cartesian product $I_1 \times I_2$ where $I_1 \subset \mathbb{R}^k$, $I_2 \subset \mathbb{R}^{m-k}$;

ii) $\forall x \in J, \quad W_x \subset W$.

Proof. We "complete" the map f by defining:

$$H : \begin{cases} \mathbb{R}^m \times \mathbb{R}^d & \to \mathbb{R}^m \times \mathbb{R}^d \\ (\mathbf{t}, x) & \mapsto (f_x(\mathbf{t}), x). \end{cases}$$

The function H is $\mathcal{C}_1$, and its derivative can be written in block form:

$$DH_{(\mathbf{t},x)} = \begin{pmatrix} (Df_x)_{\mathbf{t}} & \star \\ 0 & I_k \end{pmatrix}.$$

Since $(Df_{x_0})_{\mathbf{t}_0}$ is invertible, $(DH)_{\mathbf{t}_0,x_0}$ is invertible. We apply the local inversion theorem to H: there exist open sets $\mathcal{U}_0$, $\mathcal{V}_0$ such that H maps $\mathcal{U}_0$ to $\mathcal{V}_0$ diffeomorphically. In order to satisfy the properties i) and ii), we restrict H two times. First we define $\mathcal{U}_1 = \mathcal{U}_0 \cap (W \times \mathbb{R}^d)$, and $\mathcal{V}_1 = H(\mathcal{U}_1)$. Since $\mathcal{V}_1$ is open it contains a product set $\mathcal{V} = I_1 \times I_2 \times J$, and we let $\mathcal{U} = H^{-1}(\mathcal{V})$. For any $(y, x) \in I \times J$, define $g_x(y)$ the first component of $H^{-1}(y, x)$: composing by H, we see that $f_x(g_x(y)) = y$.

The set $W_x = \{\mathbf{t} \in \mathbb{R}^m; (\mathbf{t}, x) \in \mathcal{U}\}$ is open, and included in W. Since f_x maps W_x to I, g_x is its inverse and both are continuous, so $\tilde{f}_x$ is a diffeomorphism. □

Lemma 6.3. *Let T be a continuous random variable in $\mathbb{R}^m$, with density h_T. Let $d \leqslant m$, and let ϕ be a $\mathcal{C}^1$ map from $\mathbb{R}^m \times \mathbb{R}^d$ to $\mathbb{R}^d$:*

$$\phi : (\mathbf{t}, x) \mapsto \phi_x(\mathbf{t}).$$

Suppose that, for some $x_0, \mathbf{t}_0$, $(D\phi_{x_0})_{\mathbf{t}_0} : \mathbb{R}^m \to \mathbb{R}^d$ has full rank d. Suppose additionally that h_T is bounded below by $c_0 > 0$ on a neighborhood of $\mathbf{t}_0$.

Then there exist a constant $c > 0$, a neighborhood J of x_0 and a neighborhood I_1 of $\phi_{x_0}(\mathbf{t}_0)$ such that:

$$\forall x \in J, \quad \mathbb{P}\left[\phi(T, x) \in \cdot\right] \geqslant c\lambda_{\mathbb{R}^d}(\cdot \cap I_1). \tag{33}$$

In other words, $\phi(T, x)$ has an absolutely continuous part with respect to the Lebesgue measure.

Proof. We know that $(D\phi_{x_0})_{\mathbf{t}_0}$ has rank d. Without loss of generality, we suppose that the first d columns are independent. In other words, writing $\mathbf{t} = (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^d \times \mathbb{R}^{m-d}$, we suppose that the derivative of $\psi_{x,\mathbf{v}} : \mathbf{u} \mapsto \phi_{x_0}(\mathbf{u}, \mathbf{v})$ is invertible in $\mathbf{u}_0$ for $\mathbf{v} = \mathbf{v}_0$.

Once more, we "complete" ϕ and define:

$$f_x : \begin{cases} \mathbb{R}^d \times \mathbb{R}^{m-d} & \to \mathbb{R}^d \times \mathbb{R}^{m-d} \\ (\mathbf{u}, \mathbf{v}) & \mapsto (\phi_x(\mathbf{u}, \mathbf{v}), \mathbf{v}). \end{cases}$$

By Lemma 6.2 applied with $k = d$, we can find $I_1 \subset \mathbb{R}^d$, $I_2 \subset \mathbb{R}^{m-d}$, $J \subset \mathbb{R}^d$ and $(W_x)_{x \in J} \subset \mathbb{R}^m$ such that f_x maps diffeomorphically W_x to $I_1 \times I_2$. Call $\tilde{f}_x$ this diffeomorphism. By property ii) of the lemma, we can ensure that W_x is included in a given neighborhood of $\mathbf{t}_0$. Since $Df_x = \begin{pmatrix} D\psi_{x,\mathbf{v}} & \star \\ 0 & I \end{pmatrix}$, we can choose this neighborhood so that:

$$\forall x \in J, \forall \mathbf{t} \in W_x, \quad h_T(\mathbf{t}) \left|\det((Df_x)_{\mathbf{t}})\right|^{-1} \geqslant c' > 0. \tag{34}$$

for some strictly positive constant c'.

Write the random variable T as a couple (U, V), and let A be a Borel set included in I_1.

$$\begin{aligned}
\mathbb{P}\left[\phi(T, x) \in A\right] &\geqslant \mathbb{P}\left[\phi(T, x) \in A, V \in I_2\right] \\
&= \mathbb{P}\left[f_x(U, V) \in A \times I_2\right] \\
&\geqslant \mathbb{P}\left[(U, V) \in {\tilde{f}_x}^{-1}(A \times I_2)\right] \\
&= \int_{\tilde{f}_x^{-1}(A \times I_2)} h_T(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v} \\
&= \int_{\tilde{f}_x^{-1}(A \times I_2)} h_T(\mathbf{u}, \mathbf{v}) \left|\det((D\tilde{f}_x)_{\mathbf{u},\mathbf{v}})\right|^{-1} \cdot \left|\det((D\tilde{f}_x))_{\mathbf{u},\mathbf{v}}\right| d\mathbf{u} d\mathbf{v}.
\end{aligned}$$

Since $\tilde{f}_x^{-1}(A \times I_2) \subset W_x$, we may use the bound (34). Then we can change variables by defining $(\mathbf{s}, \mathbf{v}) = \tilde{f}_x(\mathbf{u}, \mathbf{v})$. We obtain:

$$\begin{aligned}\mathbb{P}\left[\phi(T,x) \in A\right] &\geqslant c' \int_{\tilde{f}_x^{-1}(A\times I_2)} \left|\det((D\tilde{f}_x))_{\mathbf{u},\mathbf{v}}\right| d\mathbf{u}d\mathbf{v} \\ &= c' \int_{A\times I_2} d\mathbf{s}d\mathbf{v} \\ &\geqslant c'\lambda_{\mathbb{R}^d}(A)\lambda_{\mathbb{R}^{m-d}}(I_2).\end{aligned}$$

Therefore (33) holds with $c = c'\lambda_{\mathbb{R}^{m-d}}(I_2)$. □

6.3 A slightly weaker global condition

Proposition 6.4 (Regularity at jump times — weak form)**.** *Let x_0 be a point in M, and $(\mathbf{i}, \mathbf{u})$ an adapted sequence in $\mathbb{T}_m$, such that $\min_{0\leqslant i\leqslant m} u_i > 0$.*

If $\mathbf{v} \mapsto \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x_0)$ is a submersion at $\mathbf{u}$, then there exists $\mathcal{U}_0$ a neighborhood of x_0, $\mathcal{V}_0$ a neighborhood of $\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{u}}(x_0)$ and a constant $c > 0$ such that:

$$\forall x \in \mathcal{U}_0, \quad \mathbb{P}_{x,i_0}\left[\tilde{Z}_m \in \cdot \times \{i_m\}\right] \geqslant c\lambda_{\mathbb{R}^d}(\cdot \cap \mathcal{V}_0), \tag{35}$$

where i_0 and i_m are the first and last elements of $\mathbf{i}$.

Remark 6.5. *This result is a weaker form of Theorem 4.1:*

- *the hypothesis is stronger — the sequence $(\mathbf{i}, \mathbf{u})$ must be adapted with strictly positive terms;*
- *the conclusion is weaker — it lacks uniformity for the discrete component, for the starting point and the final point.*

Proof. Recall $(U_i)_{i\geqslant 1}$ is the sequence of interarrival times of a homogeneous Poisson process. Let $\mathcal{F}_{Poi}$ be the sigma field generated by $(U_i)_{i\geqslant 1}$. Set $\mathbf{U} = (U_1, \ldots, U_m)$ and $\tilde{\mathbf{Y}} = (\tilde{Y}_0, \ldots, \tilde{Y}_m)$. By continuity, there exists a neighborhood $\mathcal{U}_0$ of x_0, and numbers $\delta_1, \delta_2 > 0$ such that $p(x, \mathbf{i}, \mathbf{v}) \geqslant \delta_2$ for all $x \in \mathcal{U}_0$ and $\mathbf{v} \in \mathbb{R}^m$ such that $\|\mathbf{v} - \mathbf{u}\| = \max_{1\leqslant i\leqslant m} |v_i - u_i| \leqslant \delta_1$. Therefore

$$\begin{aligned}\mathbb{P}_{x,i_0}\left[\tilde{X}_m \in \cdot,\, \tilde{Y}_m = i_m\right] &\geqslant \mathbb{P}\left[\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{U}}(x) \in \cdot,\ \tilde{\mathbf{Y}} = \mathbf{i},\ \|\mathbf{U} - \mathbf{u}\| \leqslant \delta_1\right] \\ &= \mathbb{E}\left[\mathbb{P}\left[\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{U}}(x) \in \cdot,\ \tilde{\mathbf{Y}} = \mathbf{i},\ \|\mathbf{U} - \mathbf{u}\| \leqslant \delta_1 \middle| \mathcal{F}_{Poi}\right]\right] \\ &\geqslant \delta_2 \mathbb{P}\left[\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{U}}(x) \in \cdot,\ \|\mathbf{U} - \mathbf{u}\| \leqslant \delta_1\right] \\ &= \delta_2\delta_3 \mathbb{P}\left[\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{U}}(x) \in \cdot \middle| \|\mathbf{U} - \mathbf{u}\| \leqslant \delta_1\right] \\ &= \delta_2\delta_3 \mathbb{P}\left[\mathbf{\Phi}^{\mathbf{i}}_{\mathbf{T}}(x) \in \cdot\right]\end{aligned}$$

where $\delta_3 = \mathbb{P}\left[\|\mathbf{U} - u\| \leqslant \delta_1\right] > 0$ and $\mathbf{T} = (T_1, \ldots, T_m)$ is a vector of independent random variables such that for each i, the distribution of T_i is given by

$$\mathbb{P}\left[T_i < t\right] = \mathbb{P}\left[U_i < t \,\middle|\, |U_i - u_i| \leqslant \delta_1\right].$$

On $[u_i - \delta_1, u_i + \delta_1]$ this is $\frac{e^{\lambda\delta_1} - e^{-\lambda(t-u_i)}}{e^{\lambda\delta_1} - e^{-\lambda\delta_1}}$ so T_i has the density

$$f_{T_i}(t) = \mathbb{1}_{[u_i-\delta_1, u_i+\delta_1]}(t) \frac{\lambda e^{-\lambda(t-u_i)}}{e^{\lambda\delta_1} - e^{-\lambda\delta_1}} \tag{36}$$

which is continuous at the point u_i.

Lemma 6.3 then applies, yielding (17), with $\mathcal{U}_0$ and $\mathcal{V}_0$ given by J and I_1 of Lemma 6.3. □

6.4 Gaining uniformity

6.4.1 Uniformity at the beginning

The uniformity on the discrete component follow from two main ideas:

- use the irreducibility and aperiodicity to move the discrete component,
- use the finite speed given by compactness to show that this can be done without moving too much.

The vector fields F^i are continuous and the space is compact, so the speed of the process is bounded by a constant C_{sp}.

Definition 6.6 (Shrinking)**.** *For any open set $\mathcal{U}$ and any $t > 0$, define $\mathcal{U}_t$ the shrunk set:*

$$\mathcal{U}_t = \{x \in A, d(x, \mathcal{U}^c) > C_{sp}t\} \tag{37}$$

This set is open, and non empty for $0 < t < t(\mathcal{U})$. If $x \in \mathcal{U}_t$, then $\mathbb{P}_{x,i}[X_t \in \mathcal{U}] = 1$.

Lemma 6.7 (Uniformity at the beginning)**.** *Let $\mathcal{U}$ be a non empty open set. There exist $0 < \varepsilon_1 < \varepsilon_2$, an integer m_b, an open set $\mathcal{U}' \subset \mathcal{U}$ and a constant c such that:*

$$\forall x \in \mathcal{U}', \forall i, j, \qquad \mathbb{P}_{x,i}[\forall t \in [\varepsilon_1, \varepsilon_2], Z_t \in \mathcal{U} \times \{j\}] \geqslant c,$$
$$\forall x \in \mathcal{U}', \forall i, j \qquad \mathbb{P}_{x,i}\left[\tilde{Z}_{m_b} \in \mathcal{U} \times \{j\}\right] \geqslant c.$$

Proof. Let $\varepsilon_2 < t(\mathcal{U})$, $\mathcal{U}' = \mathcal{U}_{\varepsilon_2}$ and $\varepsilon_1 = \varepsilon_2/2$. There is a positive probability that between $t = 0$ and $t = \varepsilon_1$, the index jumps from i to j, and does not jump again before time $t = \varepsilon_2$; the fact that $X_t \in \mathcal{U}$ is guaranteed by the definition of $\mathcal{U}'$. The second result is similar; if all jump rates are positive, we can even choose $m_b = 1$. □

6.4.2 Uniformity at the end

Lemma 6.8 (Gain of discrete uniformity)**.** *If $\mathcal{V}$ is an open set and $i \in E$, then there exist c', $\mathcal{V}'$, t_e and m_e such that, if $\mu \geqslant c(\lambda_{\mathcal{V}} \times \delta_i)$,*

$$\mu P_{t_e} \geqslant c'(\lambda_{\mathcal{V}' \times E}),$$
$$\mu \tilde{P}^{m_e} \geqslant c'(\lambda_{\mathcal{V}' \times E}).$$

The proof will use the following result:

Lemma 6.9 (Propagation of absolute continuity)**.** *There is a constant C_{div} that only depends on the set M and the vector fields F^i such that, for all $\mathcal{V}$, i,*

$$(\lambda_{\mathcal{V}} \times \delta_i)K_t \geqslant e^{-C_{div}t}(\lambda_{\mathcal{V}_t} \times \delta_i),$$

where $\mathcal{V}_t$ is the shrunk set defined in (37) *and $K_t f(x, i) = \Phi_t^i f(x, i)$.*

Proof. Since Φ_t^i is a diffeomorphism from $\mathbb{R}^d$ to itself, for any positive map f on $\mathbb{R}^d \times E$ we get by change of variables:

$$\int f(\Phi_t^i(x), i) d\lambda_{\mathcal{V}}(x) = \int f(x, i) \left|D\Phi_t^i\right|^{-1} d\lambda_{\Phi_t^i(\mathcal{V})}(x).$$

If we let $h(t) = |D\Phi_t^i|$, one of the classical interpretation of the divergence operator (see e.g. [21], Proposition 16.33) yields $h'(t) = h(t) \operatorname{div} F^i(\Phi_t^i(x))$. By compactness,

$$\exists C_{div}, \forall x \in M, \forall i, \quad \left|\operatorname{div} F^i(x)\right| \leqslant C_{div}.$$

Therefore $h(t)^{-1} \geqslant \exp(-C_{div}t)$.

Since by definition of the shrunk set, $\Phi_t^i(\mathcal{V}) \supset \mathcal{V}_t$,

$$\int f(\Phi_t^i(x), i) d\lambda_{\mathcal{V}}(x) \geqslant \exp(-C_{div}t) \int f(x, i) d\lambda_{\mathcal{V}_t}(x),$$

and Lemma 6.9 follows. □

Proof of Lemma 6.8. Fix a point $x \in \mathcal{V}$. Since the matrix $Q(x) = Q(x, i, j)$ is irreducible and aperiodic, there exists an integer m such that for all i, j, there exists a sequence $\mathbf{i}(i, j) = (i_0 = i, i_1(i, j), \ldots, i_{m-1}(i, j), i_m(i, j) = j)$ that satisfies $\prod_{l=1}^m Q(x, i_{l-1}(i, j), i_l(i, j)) > 0$. Without loss of generality (since we can always replace $\mathcal{V}$ by a smaller set) we suppose that

$$\forall x \in \mathcal{V}, \forall i, j, \forall l, \quad Q(x, i_{l-1}(i, j), i_l(i, j)) \geqslant c_Q > 0. \tag{38}$$

Fix i and j. From (7) we can rewrite P_t as:

$$P_t = \sum_{n \geqslant 0} \lambda^n e^{-\lambda t} \int_{\{\mathbf{u} \in \mathbb{R}^n : \sum_{i=1}^n u_i < t\}} \left(K_{u_1} Q K_{u_2} Q \cdots K_{u_n} Q K_{t - \sum u_i}\right) du_1 \ldots du_n.$$

Therefore:

$$(\lambda_{\mathcal{V}} \otimes \delta_i) P_t \geqslant \lambda^m e^{-\lambda t} \int_{\mathbf{u} \in \mathbb{R}^m : \sum u_i < t} (\lambda_{\mathcal{V}} \otimes \delta_i) K_{u_1} Q \cdots K_{u_m} Q K_{t - \sum u_i} du_1 \cdots du_n.$$

By Lemma 6.9 and the lower bound (38),

$$\begin{aligned}(\lambda_{\mathcal{V}} \otimes \delta_i) K_{u_1} Q &\geqslant e^{-C_{div} u_1} (\lambda_{\mathcal{V}_{u_1}} \otimes \delta_i) Q \\ &\geqslant c_Q e^{-C_{div} u_1} (\lambda_{\mathcal{V}_{u_1}} \otimes \delta_{i_1}).\end{aligned}$$

Repeating these two lower bounds m times yields:

$$\begin{aligned}(\lambda_{\mathcal{V}} \otimes \delta_i) P_t &\geqslant \lambda^m e^{-\lambda t} e^{-C_{div} t} c_Q^m \left(\int_{\mathbf{u} \in \mathbb{R}^m : \sum u_i < t} du_1 \cdots du_n \right) \lambda_{\mathcal{V}_t} \otimes \delta_j \\ &\geqslant \frac{(\lambda c_Q t)^m}{m!} e^{-(\lambda + C_{div})t} \lambda_{\mathcal{V}_t} \otimes \delta_j.\end{aligned}$$

Since the measures $\lambda_{\mathcal{V}_t} \otimes \delta_j$ are mutually singular for different indices j, this implies that

$$(\lambda_{\mathcal{V}} \otimes \delta_i) P_t \geqslant c(\lambda, t, c_Q, m) \lambda_{\mathcal{V}_t \times E}.$$

For t small enough, $\mathcal{V}_t = \mathcal{V}'$ is non empty, and the first part of the lemma follows.

The statement for $\tilde{P}^m$ is proved similarly, starting from the bound

$$\tilde{P}^m \geqslant \int_{\{\mathbf{u} \in \mathbb{R}^m : \sum_{i=1}^m u_i < t\}} (K_{u_1} Q K_{u_2} Q \cdots K_{u_n} Q) du_1 \ldots du_m,$$

written for a t small enough so that $\mathcal{V}_t$ is non empty. □

6.4.3 Proof of Theorem 4.1

The hypothesis gives the existence of $(\mathbf{i}, \mathbf{u})$ such that $\mathbf{v} \mapsto \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}(x_0)$ is a submersion at $\mathbf{u}$, or in other words that the family $\tilde{\mathcal{C}}(\mathbf{i}, \mathbf{u})$ defined by (31) has full rank. If $(\mathbf{i}, \mathbf{u})$ is not adapted to x_0, by the irreducibility hypothesis, there exists an m and a sequence $(\mathbf{i}', \mathbf{u}') \in \mathbb{T}_m$ such that $(\mathbf{i}', \mathbf{u}')$ is adapted and describes the same trajectory (just add instantaneous transitions where it is needed). The new family $\tilde{\mathcal{C}}(\mathbf{i}', \mathbf{u}')$ contains all vectors from $\tilde{\mathcal{C}}(\mathbf{i}, \mathbf{u})$, so

$$\operatorname{rank}(\tilde{\mathcal{C}}(\mathbf{i}', \mathbf{u}')) \geqslant \operatorname{rank}(\tilde{\mathcal{C}}(\mathbf{i}, \mathbf{u})).$$

Now, for any m, the mapping: $(\mathbf{i}, \mathbf{u}) \mapsto \tilde{\mathcal{C}}(\mathbf{i}, \mathbf{u})$ from $\mathbb{T}_m$ to $(\mathbb{R}^d)^m$ is continuous. Since the rank is a lower semicontinuous function, the mapping

$$\begin{aligned} \mathbb{T}_m &\to \mathbb{N} \\ (\mathbf{i}, \mathbf{u}) &\mapsto \operatorname{rank}\Big(\tilde{\mathcal{C}}(\mathbf{i}, \mathbf{u})\Big) \end{aligned}$$

is lower semi-continuous. Since being adapted is an open condition, there exists a sequence $(\mathbf{i}'', \mathbf{u}'') \in \mathbb{T}_m$ such that every component of $\mathbf{u}''$ is strictly positive, and $\operatorname{rank}\Big(\tilde{\mathcal{C}}(\mathbf{i}'', \mathbf{u}'')\Big) \geqslant \operatorname{rank}\Big(\tilde{\mathcal{C}}(\mathbf{i}', \mathbf{u}')\Big)$.

In other words, if the submersion hypothesis of Theorem 4.1 holds, then the stronger hypothesis of Proposition 6.4 holds for a (possibly longer) adapted sequence with non-zero terms.

By Proposition 6.4, there exists $\mathcal{U}$, $\mathcal{V}$ and c such that

$$\forall x \in \mathcal{U}, \quad \mathbb{P}_{x,i_0}\left[\tilde{Z}_m \in \cdot \times \{i_m\}\right] \geqslant c\lambda_{\mathcal{V}}(\cdot).$$

Using Lemma 6.7 to gain uniformity at the beginning, we get the existence of $m' = m_b + m$, $\mathcal{U}'$ and c' such that:

$$\forall x \in \mathcal{U}', \forall i \in E, \quad \mathbb{P}_{x,i}\left[\tilde{Z}_{m'} \in \cdot \times \{i_m\}\right] \geqslant c'\lambda_{\mathcal{V}}(\cdot),$$

or in other words:

$$\forall x \in \mathcal{U}', \forall i \in E, \quad (\delta_{x,i})\tilde{P}^{m'} \geqslant c'\lambda_{\mathcal{V}} \otimes \delta_{i_m}.$$

Finally we apply Lemma 6.8 to $i = i_m$ and the measure $\mu = \delta_{x,i}\tilde{P}^{m'}$ to get uniformity at the end: for $m'' = m' + m_e$,

$$\forall x \in \mathcal{U}', \forall i \in E, \quad (\delta_{x,i})\tilde{P}^{m''} \geqslant c''\lambda_{\mathcal{V}' \times E},$$

which is exactly the conclusion of Theorem 4.1.

6.5 Absolute continuity at fixed time

The hypothesis of Theorem 4.2 is that $\psi : \mathbf{v} \mapsto \Phi^i_{t-\sum v_i} \circ \mathbf{\Phi}^{\mathbf{i}}_{\mathbf{v}}$ has full rank at $\mathbf{u}$. Reasoning as in Section 6.1, we can compute the derivatives with respect to the v_i, and write the rank condition at the initial point x_0: the submersion hypothesis holds if and only if the family

$$\begin{aligned} \mathcal{C}(\mathbf{i}, \mathbf{u}) = \Big\{ &\Big(F^{i_0} - \mathbf{\Phi}^{\star}_m F^{i_m}\Big)(x_0), \Big(\mathbf{\Phi}^{\star}_1 F^{i_1} - \mathbf{\Phi}^{\star}_m F^{i_m}\Big)(x_0), \\ &\cdots, \Big(\mathbf{\Phi}^{\star}_{m-1} F^{i_{m-1}} - \mathbf{\Phi}^{\star}_m F^{i_m}\Big)(x_0) \Big\} \end{aligned} \tag{39}$$

has full rank.

Let us now turn to the proof of (18),

By the same continuity arguments as above, we suppose without loss of generality that the sequence $(\mathbf{i}, \mathbf{u})$ is adapted to x_0 and that all elements of $\mathbf{u}$ are positive. Moreover, there exist $\mathcal{U}_0$, δ_1 and δ_2 such that, if $x \in \mathcal{U}_0$ and $\mathbf{v} \in \mathbb{R}^m$ satisfies $\|\mathbf{v} - \mathbf{u}\| \leqslant \delta_1$, then $\sum v_i < t_0$ and $p(x, \mathbf{v}, \mathbf{i}) \geqslant \delta_2$.

Define two events

$$A = \text{“the process jumps exactly } m \text{ times before time } t_0\text{”} \cap \{\|\mathbf{U} - \mathbf{u}\| \leqslant \delta_1\},$$
$$B = \left\{\tilde{\mathbf{Y}} = \mathbf{i}\right\}$$

The event A is $\mathcal{F}_{Poi}$-measurable. By definition of δ_1, δ_2,

$$\begin{aligned}
\mathbb{1}_A \mathbb{P}_{x,i_0}\left[B|\mathcal{F}_{Poi}\right] &\geqslant \delta_2 \mathbb{1}_A \\
&\geqslant \delta_2 \mathbb{1}_{\{U_{k+1} > t_0\}} \mathbb{1}_{\{\|\mathbf{U}-\mathbf{u}\| \leqslant \delta_1\}} \\
&\geqslant \delta_2 e^{-\lambda t_0} \mathbb{1}_{\{\|\mathbf{U}-\mathbf{u}\| \leqslant \delta_1\}}.
\end{aligned}$$

On the event B, $Z_{t_0} = (\psi(\mathbf{U}), i_m)$, so:

$$\begin{aligned}
\mathbb{P}_{x,i_0}\left[Z_{t_0} \in \cdot \times \{i_m\}\right] &\geqslant \mathbb{P}_{x,i_0}\left[A \cap B \cap (Z_{t_0} \in \cdot \times \{i_m\})\right] \\
&= \mathbb{P}_{x,i_0}\left[A \cap B \cap (\psi(\mathbf{U}) \in \cdot)\right] \\
&= \mathbb{E}\left[\mathbb{P}_{x,i_0}\left[B|\mathcal{F}_{Poi}\right] \mathbb{1}_A \mathbb{1}_{\psi(\mathbf{U}) \in \cdot}\right] \\
&\geqslant \delta_2 e^{-\lambda t_0} \mathbb{P}\left[\{\|\mathbf{U} - \mathbf{u}\| \leqslant \delta_1\} \cap \psi(\mathbf{U}) \in \cdot\right] \\
&\geqslant \delta_2 \delta_3 e^{-\lambda t_0} \mathbb{P}\left[\psi(\mathbf{U}) \in \cdot \middle| \|\mathbf{U} - \mathbf{u}\| \leqslant \delta_1\right].
\end{aligned}$$

where $\delta_3 = \mathbb{P}\left[\|\mathbf{U} - \mathbf{u}\| \leqslant \delta_1\right]$. The reasoning leading to Equation (36) still applies. Thanks to Lemma 6.3, this implies (18), but only with $i = i_0$, $j = i_m$ and $\varepsilon = 0$.

To prove the general form of (18) with the additional freedom in the choice of i, j and t, we first use Lemma 6.7 to find a neighborhood $\mathcal{U}_0'$ of x_0, and three constants $0 < \varepsilon_1 < \varepsilon_2$ and $c > 0$ such that:

$$\forall x \in \mathcal{U}_0', \quad \mathbb{P}_{x,i}\left[\forall t \in [\varepsilon_1, \varepsilon_2], Z_t \in \mathcal{U}_0 \times \{i_0\}\right] \geqslant c.$$

Let $t_0' = t_0 - \varepsilon_1$ and $\varepsilon = \varepsilon_2 - \varepsilon_1$, so that $[t_0', t_0' + \varepsilon] = [t_0 + \varepsilon_1, t_0 + \varepsilon_2]$. Then, for any $x \in \mathcal{U}_0'$, and any $t \in [t_0', t_0' + \varepsilon]$,

$$\begin{aligned}
\mathbb{P}_{x,i}\left[X_t \in \cdot\right] &\geqslant \mathbb{E}_{x,i}\left[\mathbb{1}_{\{Z_{t-t_0} \in \mathcal{U}_0 \times \{i_1\}\}} \mathbb{P}_{Z_{t-t_0}}\left[X_{t_0} \in \cdot\right]\right] \\
&\geqslant c' \lambda_{\mathbb{R}^m}(\cdot \cap \mathcal{V}_0).
\end{aligned}$$

An application of Lemma 6.8 proves that we can also gain uniformity at the end; this concludes the proof of Theorem 4.2.

7 Constructive proofs for the local criteria

7.1 Regularity at jump times

To prove the local criteria, we show that they imply the global ones for appropriate (and small) times $u_1, \dots u_m$. We introduce some additional notation for some families of vector fields.

Definition 7.1. *The round letters $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ will denote families of vector fields on M. For a family $\mathcal{F}$, $\mathcal{F}_x$ is the corresponding family of tangent vectors at x.*

If $\mathbf{i} = (i_0, \dots, i_m)$ is a sequence of indices and $\mathbf{u} = \mathbf{u}(t) = (u_1(t), \dots u_m(t))$ is a sequence of "time" functions, we denote by $\mathcal{F}_{\mathbf{i},\mathbf{u}} = \mathcal{F}_{\mathbf{i},\mathbf{u}}(t)$ the family of vector fields:

$$\left\{F^{i_0}, \mathbf{\Phi}_1^\star F^{i_1}, \dots \mathbf{\Phi}_m^\star F^{i_m}\right\}. \tag{40}$$

This family depends on t via the $\mathbf{\Phi}_k^\star$ (see (32)).

We begin by a simple case where there are just two vector fields, F^1 and F^2, and we want regularity at a jump time, starting from (say) $(x, 1)$. To simplify matters further, suppose that the dimension d is two.

In the simplest case, $F^1(x)$ and $F^2(x)$ span the tangent plane $\mathbb{R}^2$. Then, for t small enough, these vectors "stay independent" along the flow of X_2: $F^2(x)$ and $(\Phi_t^{2,\star} F^1)(x)$ are independent. So the global condition holds for t small enough.

To understand where Lie brackets enter the picture, let us first recall that they appear as a Lie derivative that describes how X changes when pulled back by the flow of F^i for a small time: at any given point x,

$$\lim_{t\to 0} \Phi_t^{i,\star} X(x) - X(x) - t[F^i, X](x) = 0.$$

Staying at a formal level for the time being, let us write this as:

$$\Phi_t^{i,\star} X = X - t[F^i, X](x) + o(t). \tag{41}$$

Suppose now that F^1 and F^2 are collinear at x, but that $F^1(x)$ and $[F^1, F^2](x)$ span $\mathbb{R}^2$. We have just seen that:

$$\Phi_t^{2,\star}(F^1) = F^1 + t[F^2, F^1] + o(t).$$

Let $\mathbf{u}(t) = (t, t)$ and $\mathbf{i} = (1, 2, 1)$, and look at $\mathcal{F}_{\mathbf{i},\mathbf{u}}(t)$. If the "$o(1)$" terms behave as expected,

$$\begin{aligned}
\mathcal{F}_{\mathbf{i},\mathbf{u}}(t) &= (F^1,\ \Phi_t^{1,\star}(F^2),\ \Phi_t^{1,\star}\Phi_t^{2,\star}(F^1)) \\
&= \Big(F^1,\ F^2 + t[F^1, F^2] + o(t),\ F^1 + t[F^2, F^1] + o(t) + t[F^1, F^1] + t^2[F^1, [F^2, F^1]] + o(t^2)\Big) \\
&= \Big(F^1,\ F^2 + t[F^1, F^2] + o(t),\ F^1 + t[F^2, F^1] + o(t)\Big).
\end{aligned}$$

By hypothesis, $\operatorname{rank}\big(F^1(x), [F^2, F^1](x)\big) = 2$. Therefore, for t small enough, the lower-semicontinuity of the rank ensures:

$$\begin{aligned}
\operatorname{rank}\left(\mathcal{F}_{\mathbf{i},\mathbf{u}}(t)\right) &= \operatorname{rank}\Big(F^1,\ F^2 + t[F^1, F^2] + o(t),\ t[F^2, F^1] + o(t)\Big) \\
&= \operatorname{rank}\Big(F_1,\ F_2 + t[F^1, F^2] + o(t),\ [F^2, F^1] + o(1)\Big) \\
&\geqslant \operatorname{rank}(F^1,\ F^2,\ [F^2, F^1]) \\
&= 2,
\end{aligned}$$

and the global condition holds.

To treat the general case, let us first define the $o(1)$ notation. For any smooth vector field $X = \sum_i X^i(x)\frac{\partial}{\partial x_i}$, let

$$\|X\|_k = \max_{\alpha, |\alpha| \leqslant k} \max_{x\in M} \max_i \left|\partial^\alpha X^i(x)\right|.$$

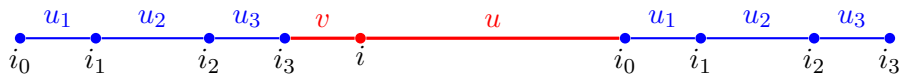


Figure 5: The Hanoï construction
The sequence $(\tilde{\mathbf{i}}; \tilde{\mathbf{u}})$ is obtained by inserting the sequence $(i; (u, v))$, depicted in red, in-between two copies of the original sequence $(\mathbf{i}; \mathbf{u})$. The "middle" time u is much larger than all the other times.

If $X(t)$ is a family of vector fields depending on the parameter t, we write $X(t) = o(1)$ if

$$\forall k \geqslant 0, \quad \lim_{t \to 0} \|X(t)\|_k = 0. \tag{42}$$

The previous case shows two main ingredients in the proof:

- to introduce the brackets, we have to alternate between flows,
- the method works because we can express various vectors and Lie brackets as (approximate) linear combinations of vectors in $\mathcal{F}_{\mathbf{i},\mathbf{u}}(t)$, for good choices of $\mathbf{u}$ and $\mathbf{i}$.

Let us abstract the second ingredient in a definition.

Definition 7.2. *Let $\mathcal{G} = \{G_1, \dots G_n\}$ be a fixed family of vector fields and $\mathcal{H}(t)$ a family depending on t. If there exist continuous functions $\lambda_{ij} : (0, \infty) \to \mathbb{R}$, and vector fields $R_i(t)$, such that:*

$$\forall i \leqslant n, \quad G_i = \sum \lambda_{ij}(t) H_j(t) + R_i(t),$$

and $R_i(t) = o(1)$, we say that $\mathcal{H}(t)$ (asymptotically) generates $\mathcal{G}$.

Remark 7.3. *We allow the $\lambda_{ij}(t)$ to blow up when $t \to 0$. Because of this, even if the $H_j(0)$ are defined, G_i does not necessary lie in their linear span. For example, if $H_1(t) = (1, 0)$ and $H_2(t) = (0, t)$, $\mathcal{H}(t)$ asymptotically generates any family of constant vector fields in $\mathbb{R}^2$, even if $H_2(0) = (0, 0)$ degenerates.*

Lemma 7.4. *If $\mathcal{H}(t)$ generates $\mathcal{G}$, then for any point x, there is a t_x such that:*

$$\forall t < t_x, \quad \operatorname{rank}(\mathcal{H}(t)_x) \geqslant \operatorname{rank}(\mathcal{G}_x).$$

This result will be proved below. The last ingredient in the proof is to introduce different time scales in the alternation between vector fields. Let us write $u(t) \ll v(t)$ if $u(t) = o(v(t))$ when t goes to zero.

Lemma 7.5 (Towers of Hanoï)**.** *Suppose that, for some $\mathbf{i} = (i_0, \dots i_m)$ and $\mathbf{u}(t) = (u_1(t), \dots u_m(t))$, $\mathcal{F}_{\mathbf{i},\mathbf{u}}(t)$ generates $\mathcal{G}$. Suppose that for all j, $u_j(t)$ takes positive values and $\lim_{t \to 0} u_j(t) = 0$. Choose two functions $u(t)$ and $v(t)$ such that $v(t) \ll u(t)$, $u(t) \ll 1$ and $u_j(t) \ll u(t)$, for all j. Define $\tilde{\mathbf{u}}(t)$ and $\tilde{\mathbf{i}}$ by concatenation (see Figure 5):*

$$\begin{aligned} \tilde{\mathbf{u}}(t) &= (u_1(t), u_2(t), \dots, u_m(t),\ v(t), u(t),\ u_1(t), \dots, u_m(t)); \\ \tilde{\mathbf{i}} &= (i_0, i_1, \dots i_m, i, i_0, i_1, \dots i_m). \end{aligned} \tag{43}$$

Then $\mathcal{F}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}(t)$ generates $\mathcal{G} \cup \{F^i\} \cup \{[F^i, G], G \in \mathcal{G}\}$.

The name comes from the fact that, in analogy with the towers of Hanoï, in order to gain brackets by F^i (the lower disk), we have to move according to $\mathbf{i}, \mathbf{u}$ (move all the upper disks), then move the lower disk, then move the upper disks once again.

Once these lemmas are known, Theorem 4.4 follows quite easily.

Proof of Theorem 4.4. Let us reason by induction. Starting from the empty family, we can construct thanks to Lemma 7.5 a $\mathbf{u}(t)$ and an $\mathbf{i}$ such that $\mathcal{F}_{\mathbf{i},\mathbf{u}}(t)$ generates any number of F^i and iterated brackets $[F^{i_1}, [F^{i_2}, \dots F^{i_k}]]$. If $\mathbb{R}^d$ is generated by these brackets, Lemma 7.4 shows that $\mathcal{F}_{\mathbf{i},\mathbf{u}}(t)$ has full rank for some t_x. Therefore, the global condition is satisfied for the choice of times $(u_1(t_x), u_2(t_x), \dots u_m(t_x))$. □

Let us turn to the proofs of the lemmas.

Proof of Lemma 7.4. Define a family $\tilde{\mathcal{H}}(t)$ by $\tilde{H}_i = \sum \lambda_{ij}(t) H_j(t) = G_i - R_i(t)$. At every point x, every vector in $\tilde{\mathcal{H}}_x$ is a combination of vectors in $\mathcal{H}_x$. Therefore

$$\begin{aligned} \operatorname{rank}(\mathcal{H}(t)) &\geqslant \operatorname{rank}(\tilde{\mathcal{H}}(t)) \\ &\geqslant \operatorname{rank}(\mathcal{G}), \end{aligned}$$

where the second line follows from the lower semi-continuity of the rank, since for all x, $\|R_i(t)(x)\|$ converges to 0. □

To prove Lemma 7.5 we need to write down properly the fundamental relation (41) and see how "small" vector fields are affected by pullbacks and Lie brackets.

Lemma 7.6. *For any k, there exists a constant C_k that only depends on the fields F^i, such that the following holds. For any vector field X, any i and any $t \in [0,1]$,*

$$\left\| \Phi_t^{i,\star} X \right\|_k \leqslant C_k \left\| X \right\|_k, \tag{44}$$

$$\left\| \Phi_t^{i,\star} X - X \right\|_k \leqslant C_k t \left\| X \right\|_{k+1}, \tag{45}$$

$$\left\| \Phi_t^{i,\star} X - X - t[F^i, X] \right\|_k \leqslant C_k t^2 \left\| X \right\|_{k+2}. \tag{46}$$

In particular if X is smooth, the formal equation (41) *is rigorous.*

Moreover, if $X(t) = o(1)$, then $\Phi_t^{i,\star} X(t) = o(1)$ and $[F^i, X(t)] = o(1)$.

Proof. Let F be one of the F^i, and $(t,x) \mapsto \Phi_t(x)$ be its flow. Recall that the pullback $\Phi_t^\star$ acts on vector fields by:

$$(\Phi_t^\star X)(x) = (D\Phi_t)_x^{-1} X(\Phi(t,x)).$$

Since the flow is smooth, this can be written in coordinates as

$$(\Phi_t^\star X)(x) = \sum_{i,j} a_j^i(x,t) X^j(\Phi(t,x)) \frac{\partial}{\partial x_i}, \tag{47}$$

where the a_j^i are smooth functions of x and t that only depend on the vector field F. For any i and any multiindex α, we may apply ∂^α to the i^{th} coordinate; the resulting expression only involves the derivatives of X^j up to order k. This implies the control (44).

To prove (45), first apply Taylor's formula in the t variable to each coordinate in (47), at order 1:

$$\begin{aligned} &a_j^i(x,t) X^j(\Phi(t,x)) - X^j(x) \\ &\quad = \int_0^t \partial_s (a_j^i(x,s) X^j(\Phi(s,x)))ds \\ &\quad = \int_0^t (\partial_s a_j^i)(x,s) X^j(\Phi(s,x)) + a_j^i(x,s) \left(\nabla X^j(\Phi(s,x)) \cdot F(\Phi(s,x)) \right) ds. \end{aligned}$$

This expression involves the X^j and their first-order derivatives. Once more we may apply ∂^α to both sides to deduce (45). The proof of (46) is similar; the fact that the first order term is given by the Lie bracket is standard. This clearly implies (41) if X is smooth.

Finally suppose $X(t)$ satisfies $X(t) = o(1)$. For any k, $\left\|\Phi_t^{i,\star}X(t)\right\|_k \leqslant C_k \left\|X(t)\right\|_k$ by (44). When t goes to zero, $\left\|X(t)\right\|_k$ converges to zero, and so does $\left\|\Phi_t^{i,\star}X(t)\right\|_k$: in other words, $\Phi_t^{i,\star}X(t) = o(1)$. The fact that the same happens for Lie brackets follows from their expression in coordinates. □

Proof of Lemma 7.5. Recalling the composite pullback notation $\mathbf{\Phi}_k^\star$ from (32), let us define, for $0 \leqslant j \leqslant m$, $\tilde{F}_j = \mathbf{\Phi}_j^\star F^{i_j}$ so that $\mathcal{F}_{\mathbf{i},\mathbf{u}}(t) = (\tilde{F}_0, \tilde{F}_1, \dots \tilde{F}_m)$. All these quantities, as well as u, depend on t, but we drop this dependence in the notation. The "Hanoï" construction yields:

$$\begin{aligned}
\mathcal{F}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}} = (\quad & F^{i_0}, && \mathbf{\Phi}_1^\star F^{i_1}, && \dots, && \mathbf{\Phi}_m^\star F^m, \\
& && \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star} F^i, && && \\
& \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star} F^{i_0}, && \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star}\mathbf{\Phi}_1^\star F^{i_1}, && \dots, && \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star}\mathbf{\Phi}_m^\star F^{i_m}) \\
= (\quad & \tilde{F}_0, && \tilde{F}_1, && \dots, && \tilde{F}_m, \\
& && \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star} F^i, && && \\
& \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star} \tilde{F}_0, && \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star}\tilde{F}_1, && \dots, && \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star}\tilde{F}_m).
\end{aligned}$$

Consider first the middle term. Since $v = v(t)$ goes to zero, $\Phi_v^{i_m,\star}F^i = F^i + o(1)$ by (45) from Lemma 7.6. Since $u_j(t)$ goes to zero, the same argument and the fact that pullbacks of $o(1)$ terms stay $o(1)$, once more by Lemma 7.6, show that $\Phi_{u_j}^{i_{j-1},\star}(F^i + o(1)) = F^i + o(1)$. Therefore:

$$\mathbf{\Phi}_m^\star \Phi_v^{i_m,\star} F^i = F^i + o(1)$$

so that $\mathcal{F}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}(t)$ generates F^i.

Now consider $\mathcal{G} = (G_1, \dots G_L)$. Since $\mathcal{F}_{\mathbf{i},\mathbf{u}}(t)$ generates $\mathcal{G}$, and $\mathcal{F}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}(t)$ contains $\mathcal{F}_{\mathbf{i},\mathbf{u}}(t)$, $\mathcal{F}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}(t)$ generates all G_l. Let us now prove that $\mathcal{F}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}(t)$ generates $[F^i, G_l]$. The first step is to write down a relation for G_l:

$$G_l = \sum_j \lambda_{lj}(t)\tilde{F}_j(t) + R_l(t).$$

where $R_l(t) = o(1)$. Since this is an equality of vector fields, we can pull it back by $\mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star}$:

$$\mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star} G_l = \sum_j \lambda_{lj}(t)\mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star}\tilde{F}_j + \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star} R_l(t).$$

The difference between the last two equalities yields

$$\mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star} G_l - G_l = \sum_j \lambda_{lj}(t)(\mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star}\tilde{F}_j - \tilde{F}_j(t)) + \mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star} R_l(t) - R_l(t). \quad (48)$$

Since the u_j and v are negligible with respect to u, repeated applications of Lemma 7.6 show that the left hand side can be written as

$$\mathbf{\Phi}_m^\star \Phi_v^{i_m,\star}\Phi_u^{i,\star} G_l - G_l = u[F^i, G_l] + o(u).$$

Similarly the last term on the right hand side satisfies

$$\mathbf{\Phi}^\star_m \Phi_v^{i_m,\star} \Phi_u^{i,\star} R_l(t) - R_l(t) = u[F^i, R_l(t)] + o(u).$$

Plugging this back in (48) and dividing by u we get

$$[F_i, G_l] = \sum_j \frac{\lambda_{lj}(t)}{u} (\mathbf{\Phi}^\star_m \Phi_v^{i_m,\star} \Phi_u^{i,\star} \tilde{F}_j - \tilde{F}_j(t)) + o(1) + [F_i, R_l(t)].$$

Since $R_l(t) = o(1)$, the last statement of Lemma 7.6 implies that $[F^i, R_l(t)] = o(1)$. Therefore $[F^i, G_l]$ can be written as the sum of a linear combination of vector fields in $\mathcal{F}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}$, and a remainder. This shows that $\mathcal{G} \cup \{F^i\} \cup \{[F^i, G], G \in \mathcal{G}\}$ is generated by $\mathcal{F}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}$, and concludes the proof of Lemma 7.5. □

7.2 Regularity at a fixed time

Once more, we show that if the local criterion holds, then the global one holds for a good choice of indices and times. The global criterion is expressed in terms of the family described by (39). It will be easier to work with a slightly different family, namely:

$$\mathcal{G}_{\mathbf{i},\mathbf{u}} = \left(F^{i_0} - \mathbf{\Phi}^\star_1 F^{i_1}, \mathbf{\Phi}^\star_1 F^{i_1} - \mathbf{\Phi}^\star_2 F^{i_2}, \dots, \mathbf{\Phi}^\star_{m-1} F^{i_{m-1}} - \mathbf{\Phi}^\star_m F^{i_m} \right). \tag{49}$$

It is easy to see that $\mathcal{G}_{\mathbf{i},\mathbf{u}}$ and the original family span the same space at each point (the k^{th} vector in the original family is the sum of the last $m-k+1$ vectors of $\mathcal{G}_{\mathbf{i},\mathbf{u}}$). The analogue of Lemma 7.5 is the following:

Lemma 7.7 (More Towers of Hanoï)**.** *Suppose that, for some* $\mathbf{i}$ *and some time functions* $\mathbf{u}(t)$*,* $\mathcal{G}_{\mathbf{i},\mathbf{u}}(t)$ *asymptotically generates* $\mathcal{G}$*. Choose* $u(t)$*,* $v(t)$ *as in Lemma 7.5 and define* $\tilde{\mathbf{i}}$*,* $\tilde{\mathbf{u}}(t)$ *by concatenation as in* (43)*.*

Then $\mathcal{G}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}(t)$ *asymptotically generates* $\mathcal{G} \cup \{F^i - F^{i_m}, F^i - F^{i_0}\} \cup \{[F^i, G], G \in \mathcal{G}\}$*.*

Proof. For $1 \leqslant j \leqslant m$, call $\tilde{F}_j$ the j^{th} vector field in $\mathcal{G}_{\mathbf{i},\mathbf{u}}$: $\tilde{F}_j = \mathbf{\Phi}^\star_{j-1} F^{i_{j-1}} - \mathbf{\Phi}^\star_j F^{i_j}$. The family $\mathcal{G}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}$ is then

$$\begin{aligned}
\mathcal{G}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}} = (\quad & \tilde{F}_1, && \tilde{F}_2, && \dots, && \tilde{F}_m, \\
& \mathbf{\Phi}^\star_m (F^{i_m} - \Phi_v^{i_m,\star} F^i), && \mathbf{\Phi}^\star_m \Phi_v^{i_m,\star} (F^i - \Phi_u^{i,\star} F^{i_0}), && && \\
& \mathbf{\Phi}^\star_m \Phi_v^{i_m,\star} \Phi_u^{i,\star} \tilde{F}_1, && \mathbf{\Phi}^\star_m \Phi_v^{i_m,\star} \Phi_u^{i,\star} \tilde{F}_2, && \dots, && \mathbf{\Phi}^\star_m \Phi_v^{i_m,\star} \Phi_u^{i,\star} \tilde{F}_m,).
\end{aligned}$$

The two vector fields in the middle give, at zero-th order as t goes to zero, $F^{i_m} - F^i$ and $F^i - F^{i_0}$. For any element $G_l \in \mathcal{G}$, we may repeat the exact same argument as in Lemma 7.5 to generate the vectors G_l and $[F^i, G_l]$ from the family $\mathcal{G}_{\tilde{\mathbf{i}},\tilde{\mathbf{u}}}(t)$. □

With this lemma in hand, we know we can generate $F^1 - F^2$ (starting from $\mathbf{i} = (i_0) = (1)$, an empty $\mathbf{u} = ()$, and choosing $i = 2$). In all successive "enrichments" of $\mathbf{i}, \mathbf{u}$ by the Hanoï procedure, the first and last components of $\mathbf{i}$ will always be 1, therefore given enough enrichments, we generate all the $F^i - F^1$. Consequently we also get all differences: $F^i - F^j = F^i - F^1 + F^1 - F^j$. Finally, by taking the bracket by F^i, we generate $[F^i, F^j]$, and then all subsequent higher order brackets. This concludes the proof of Theorem 4.4.

Acknowledgements FM and PAZ thank MB for his kind hospitality and his coffee breaks. We acknowledge financial support from the Swiss National Foundation Grant FN 200021-138242/1 and the French ANR projects EVOL, ProbaGeo and ANR-12-JS01-0006 - PIECE. This work was mainly done while PAZ (resp. FM) held a position at the University of Burgundy (resp. Rennes).

References

[1] J.P. Aubin and A. Cellina, *Differential Inclusions: Set-Valued Maps and Viability Theory*, Springer-Verlag New York, Inc. Secaucus, NJ, USA, 1984. 3.1

[2] Y. Bakhtin and T. Hurth, *Invariant densities for dynamical systems with random switching*, Nonlinearity **25** (2012), no. 10, 2937–2952. MR 2979976 1, 1, 3.8, 3.18, 4.1, 4.1, 4.2.1, 5.1

[3] M. Benaïm, J. Hofbauer, and S. Sorin, *Stochastic approximations and differential inclusions*, I. SIAM Journal on Optimization and Control **44** (2005), 328–348. 3.10

[4] 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. 1.1

[5] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt, *On the stability of planar randomly switched systems*, Ann. Appl. Probab. **24** (2014), no. 1, 292–311. MR 3161648 1.2

[6] O. Boxma, H. Kaspi, O. Kella, and D. Perry, *On/Off Storage Systems with State-Dependent Inpout, Outpout and Swithching Rates*, Probability en the Engineering and Informational Siences **19** (2005), 1–14. 1

[7] E. Buckwar and M. G. Riedler, *An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution*, J. Math. Biol. **63** (2011), no. 6, 1051–1093. MR 2855804 (2012h:92075) 1

[8] B. Cloez and M. Hairer, *Exponential ergodicity for Markov processes with random switching*, Preprint available on arXiv:1303.6999 [math.PR]. To appear in Bernoulli, (2013). 1.2

[9] O. L. V. Costa and F. Dufour, *Stability and ergodicity of piecewise deterministic Markov processes*, SIAM J. Control Optim. **47** (2008), no. 2, 1053–1077. 2.2

[10] M. H. A. Davis, *Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models*, J. Roy. Statist. Soc. Ser. B **46** (1984), no. 3, 353–388, With discussion. MR MR790622 (87g:60062) 1

[11] ———, *Markov models and optimization*, Monographs on Statistics and Applied Probability, vol. 49, Chapman & Hall, London, 1993. 1, 2

[12] P. Diaconis and D. Freedman, *Iterated random functions*, SIAM Rev. **41** (1999), no. 1, 45–76. MR 1669737 (2000c:60102) 1

[13] M. Duflo, *Random Iterative Models*, Springer Paris, 2000. 2.2

[14] V. Dumas, F. Guillemin, and Ph. Robert, *A Markovian analysis of additive-increase multiplicative-decrease algorithms*, Adv. in Appl. Probab. **34** (2002), no. 1, 85–111. 1

[15] R. Durrett, *Stochastic calculus*, Probability and Stochastics Series, CRC Press, Boca Raton, FL, 1996, A practical introduction. MR 1398879 (97k:60148) 4.2.1

[16] H. Furstenberg, *Strict ergodicity and transformation of the torus*, Amer. J. Math. **83** (1961), 573–601. MR 0133429 (24 #A3263) 3.18

[17] C. Graham and Ph. Robert, *Interacting multi-class transmissions in large stochastic networks*, Ann. Appl. Probab. **19** (2009), no. 6, 2334–2361. 1

[18] ———, *Self-adaptive congestion control for multiclass intermittent connections in a communication network*, Queueing Syst. **69** (2011), no. 3-4, 237–257. MR 2886470 1

[19] M. Jacobsen, *Point process theory and applications*, Probability and its Applications, Birkhäuser Boston Inc., Boston, MA, 2006, Marked point and piecewise deterministic processes. MR 2189574 (2007a:60001) 1, 2

[20] R. Karmakar and I. Bose, *Graded and binary responses in stochastic gene expression*, Physical Biology **197** (2004), no. 1, 197–214. 5.2

[21] J. M. Lee, *Introduction to smooth manifolds*, second ed., Graduate Texts in Mathematics, vol. 218, Springer, New York, 2013. MR 2954043 6.4.2

[22] P. A. W Lewis and G. S. Shedler, *Simulation of nonhomogeneous poisson processes by thinning*, Naval Research Logistics Quarterly **26** (1979), no. 3, 403–413. 2

[23] T. Lindvall, *Lectures on the coupling method*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1992, A Wiley-Interscience Publication. MR 1180522 (94c:60002) 4.2.1, 4.2.2

[24] R. Mañé, *Ergodic theory and differentiable dynamics*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 8, Springer-Verlag, Berlin, 1987, Translated from the Portuguese by Silvio Levy. MR 889254 (88c:58040) 3.18

[25] 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 (2011m:60070) 1

[26] N. S Papageorgiu, *Existence Theorems for Differential Inclusions with Nonconvex Right Hand Side*, Internat. J. Math. & Sci. **9** (1986), no. 3, 459–469. 3.1

[27] O. Radulescu, A. Muller, and A. Crudu, *Théorèmes limites pour des processus de Markov à sauts. Synthèse des résultats et applications en biologie moléculaire*, Technique et Science Informatiques **26** (2007), no. 3-4, 443–469. 1, 5.2

[28] S. M. Ross, *Simulation*, second ed., Statistical Modeling and Decision Science, Academic Press Inc., San Diego, CA, 1997. MR 1433593 (99c:65013) 2

[29] G. G. Yin and C. Zhu, *Hybrid switching diffusions*, Stochastic Modelling and Applied Probability, vol. 63, Springer, New York, 2010, Properties and applications. MR 2559912 (2010i:60226) 1

Michel Benaïm, e-mail: michel.benaim(AT)unine.ch

Institut de Mathématiques, Université de Neuchâtel, 11 rue Émile Argand, 2000 Neuchâtel, Suisse.

Stéphane Le Borgne, e-mail: stephane.leborgne(AT)univ-rennes1.fr

IRMAR UMR 6625, CNRS-Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France.

Florent Malrieu, e-mail: florent.malrieu(AT)univ-tours.fr

LMPT UMR 7350, CNRS-Université de Tours, UFR Sciences et Techniques, Parc de Grandmont, 37200 Tours, France.

Pierre-André Zitt, e-mail: pierre-andre.zitt(AT)univ-mlv.fr

LAMA UMR 8050, CNRS-Université-Paris-Est-Marné-La-Vallée, 5, boulevard Descartes, Cité Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2, France.