



HAL
open science

Averaging for slow-fast piecewise deterministic Markov processes with an attractive boundary

Alexandre Genadot

► **To cite this version:**

Alexandre Genadot. Averaging for slow-fast piecewise deterministic Markov processes with an attractive boundary. 2022. hal-03531257

HAL Id: hal-03531257

<https://inria.hal.science/hal-03531257>

Preprint submitted on 18 Jan 2022

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.

Averaging for slow-fast piecewise deterministic Markov processes with an attractive boundary.

Alexandre Génadot*

Winter 2022

Abstract

In this paper, we consider the problem of averaging for a class of piecewise deterministic Markov processes (PDMP) whose dynamic is constrained by the presence of a boundary. When reaching the boundary, the process is forced to jump away from it. We assume that this boundary is attractive for the process in question in the sense that its averaged flow is not tangent to it. Our averaging result relies strongly on the existence of densities for the process, allowing us to study the average number of crossings of a smooth hypersurface by an unconstrained PDMP and to deduce from this study averaging results for constrained PDMP.

1 Introduction

In this article, we consider Piecewise Deterministic Markov Processes (PDMP in the sequel), introduced by Davis in [3] in their euclidean-mode setting. In this framework, a PDMP consists in a couple $(X(t), Y(t))_{t \geq 0}$ on $\mathbf{R}^d \times \mathbf{Y}$, where \mathbf{Y} is, for us, a finite set. The motion of the euclidean variable X is the one described by a switching ordinary differential equation, whose switches are parametrized by the mode process Y . This discrete \mathbf{Y} -valued process Y jumps between its possible states at X -dependent rates. We are interested in the interaction of the X -component with a hypersurface \mathbf{B} of \mathbf{R}^d , when the dynamic of the underlying discrete process Y is infinitely accelerated, resulting in a two-time scale process $(X_\varepsilon, Y_\varepsilon)$. The (small) real ε indicates that the time scale for Y_ε is of order ε while for X_ε , the time-scale remains of order 1. We will also be interested in the averaging of the process when the X -component is forced to jump, according to some boundary jump measure, when reaching the hypersurface \mathbf{B} considered as a boundary. We are thus in the framework of averaging for constrained Markov processes.

Averaging for unconstrained Markov process, that is, without the presence of a boundary, has been studied by several authors since decades and is well understood for a wide variety of Markov processes. For instance, see [22] for continuous time Markov chains, [17] for diffusion processes, [10] for fractional Brownian motion, [5, 6, 16] for piecewise deterministic Markov processes and [18] and references therein for a rich variety of other examples and applications. The aim, when averaging the dynamic of the process, is to obtain a process that is simpler to study, but with properties that are qualitatively close to the initial model. This explains the success of this method and why it has been, and still is, the object of so much attention. For example, see [16] for applications of this method in neuroscience.

To the best of our knowledge, averaging for constrained Markov processes, that is with the presence of a boundary, has not been the object of so much investigations, in particular in the description of the averaging measure at the boundary. There is an intrinsic difficulty in these problems which stems from the presence of two types of jumps. High-rate stochastic jumps, those of the Y component whose dynamic is accelerated, and boundary forced jumps, those of the X component. In [1], we adopted a standard approach to study a class of one dimensional and piecewise linear Markov processes with constraint. This approach is called the "penalty

*Univ. Bordeaux, CNRS, INRIA, Bordeaux INP, IMB, UMR 5251, F-33400 Talence, France.

method”, exposed in [15, Section 6.4]. This method consists in considering a penalized process jumping at fast rate when beyond the boundary rather than a process jumping instantaneously at the boundary. Then, a time change is performed in order to sufficiently slow down the dynamic of the penalized process when beyond the boundary, allowing the application of classical limit theorems for Markov processes. Let us also mention that in [6], the authors set up a general method to study Markov processes with constraints. They called this method the ”patchwork martingale problem”. It has been successfully applied in [3] to reflected diffusions.

Yet another approach is used in the present article to overcome the difficulty of handling stochastic jumps at fast rates with the additional presence of forced jumps. As far as we know, this approach is fairly original. In a first time, we circumvent the difficulty in studying the interplay between the two-time scale PDMP and the hypersurface \mathbf{B} without forced jumps. For this purpose, we use the study on the average number of crossings for piecewise smooth processes performed in [1, 2]. In these articles, Rice’s formula for the average number of crossings of a hypersurface are obtained and the so-called crossing measure at the boundary is studied more specifically in [2]. One of the main assumption in these articles is the existence of densities for the process at any time. The problem of the existence of densities for PDMP has been extensively studied in the articles [9, 21]. In our setting, a set of sufficient assumptions consists to suppose that the initial value of the process has a density, as well as the jump measure at the boundary. Once the interplay between averaging and the crossing hypersurface described, we are able to handle the presence of forced jump in the dynamic of the PDMP, that is to tackle the problem of averaging for a constrained PDMP.

The present article is organised as follows. Section 2 is devoted to the study of PDMP without boundary. We present in Section 2.1 the construction of a PDMP in the euclidean-mode setting without boundary and recall an important averaging result obtained in this context. Then, in Section 2.2, we precise what this averaging result implies for the densities of the process, when they exist. Section 3 is devoted to the description of averaging in interaction with a hypersurface. The average number of crossings is considered and convergence results for this number are presented in Section 3.1. These results are applied in Section 3.2 to obtain an averaging result for a PDMP with the presence of a fully attractive boundary. An example is considered in Section 4 as well as two possible extensions of our results. Section 5.3 is devoted to the proofs and in particular to the proof of our averaging theorem for PDMP with an attractive boundary.

2 Averaging for a class of PDMP without boundary

2.1 The model and its averaging

Notations and general assumptions. As usual, \mathbf{R} is the set of real numbers with \mathbf{R}_+^* the set of positive real numbers. The space \mathbf{R}^d , where $d \geq 2$ is an integer, is endowed with the euclidean distance d with corresponding norm $\|\cdot\|$ and associated scalar product between two points x and \tilde{x} in \mathbf{R}^d denoted by $x \cdot \tilde{x}$. \mathbf{N} is the set of positive integers. We denote by λ_d the Lebesgue measure on \mathbf{R}^d . In all what follows, $\mathbf{Y} = \{y_1, \dots, y_{|\mathbf{Y}|}\}$ is a finite set of cardinal $|\mathbf{Y}|$. The real $T > 0$ denotes a finite time horizon.

Assumption A. *The process that we consider is built from the following features. Assumptions made about these characteristics are assumed to be true throughout the article.*

A.1 *For $y \in \mathbf{Y}$, F_y is a continuously differentiable and bounded vector field on \mathbf{R}^d . Moreover, we assume that there exist constants $\kappa_1, \kappa_2 > 0$ such that*

$$\|F_y(x)\| \leq \kappa_1 + \kappa_2 \|x\|, \quad \forall x \in \mathbf{R}^d.$$

These hypotheses on the vector fields imply that for each $y \in \mathbf{Y}$, the Cauchy problem:

$$\frac{d}{dt}x(t) = F_y(x(t))$$

with initial condition $x(0) = x$ has a unique global solution generating a flow denoted by $\phi_y(x, \cdot)$.

A.2 For $(y, z) \in \mathbf{Y}^2$, $y \neq z$, the function $q_{yz} : x \in \mathbf{R}^d \mapsto q_{yz}(x) \in \mathbf{R}_+^*$ is continuously differentiable and bounded. To this family of functions is associated a $|\mathbf{Y}| \times |\mathbf{Y}|$ jump intensity matrix $Q(x)$ (also called the transition rate matrix), for each $x \in \mathbf{R}^d$, defined, for each $(y, z) \in \mathbf{Y}^2$, by

$$Q_{yz}(x) := \begin{cases} q_{yz}(x) & \text{if } z \neq y, \\ -q_y(x) & \text{if } z = y, \end{cases}$$

where $q_y(x) := \sum_{z \neq y} q_{yz}(x)$.

Description of the model. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbf{P})$ denotes a filtered probability space. We define a PDMP in the euclidean-mode setting as a càdlàg stochastic process $(X(t), Y(t))_{t \in [0, T]}$ on the state space $\mathbf{R}^d \times \mathbf{Y}$ as follows. The continuous component, X , takes values in \mathbf{R}^d and has continuous paths, while the discrete one, Y , takes values in \mathbf{Y} . The dynamic of the continuous component is defined by the family of d -dimensional vector fields F_y indexed by $y \in \mathbf{Y}$. The behaviour of the discrete component is determined by the family of rate functions q_{yz} for each $(y, z) \in \mathbf{Y}^2$, describing the jumping rate between states y and z . The sample paths of a PDMP $(X(t), Y(t))_{t \in [0, T]}$ are then constructed in the following way.

For any $(x, y) \in \mathbf{R}^d \times \mathbf{Y}$ we define the survivor function

$$S_{(x, y)}(t) := e^{-\int_0^t q_y(\phi_y(x, s)) ds}, \quad t \geq 0. \quad (1)$$

This is the survivor function of a non-negative random variable. Let $\tau_0 = \sigma_0 = 0$ and $(X(0), Y(0)) = (X_0, Y_0)$ be a $\mathbf{R}^d \times \mathbf{Y}$ -valued random variable. Assume that the process is built until time τ_{n-1} , for $n \in \mathbf{N}$. Then, we can define a random variable σ_n , the inter-jump time, satisfying

$$\mathbf{P}(\sigma_n > t | (X_{n-1}, Y_{n-1}) = (x, y)) = S_{(x, y)}(t), \quad t \geq 0.$$

We define the n th jump time by

$$\tau_n := \tau_{n-1} + \sigma_n$$

and we set

$$(X(t), Y(t)) := \begin{cases} (\phi_{Y_{n-1}}(X_{n-1}, t - \tau_{n-1}), Y_{n-1}) & \text{for } \tau_{n-1} \leq t < \tau_n, \\ (X_n, Y_n) & \text{for } t = \tau_n. \end{cases}$$

The n th post-jump location (X_n, Y_n) is a $\mathbf{R}^d \times \mathbf{Y}$ -valued random variable such that

$$X_n = \phi_{Y_{n-1}}(X_{n-1}, \tau_n - \tau_{n-1})$$

and

$$\mathbf{P}(Y(\tau_n) = z | Y(\tau_n^-) = y) = \frac{q_{yz}(X(\tau_n))}{q_y(X(\tau_n))}.$$

We built thus the trajectory up to the time horizon T .

The following theorem states that the described process is a strong Markov process and characterizes its extended generator.

Theorem 1. [3, Theorem 26.14] Suppose that Assumption A holds. There exists a filtered probability space such that a standard PDMP $(X(t), Y(t))_{t \in [0, T]}$ as constructed above is a homogeneous strong Markov process. The extended generator \mathcal{A} of the process is given by

$$\mathcal{A}f(x, y) := F_y(x) \cdot \nabla_x f(x, y) + \sum_{z \in \mathbf{Y}} q_{yz}(x)(f(x, z) - f(x, y))$$

for functions partially differentiable with respect to the variable x and belonging to the domain $D(\mathcal{A})$, which is the set of all measurable functions $f : \mathbf{R}^d \times \mathbf{Y} \rightarrow \mathbf{R}$, such that $t \mapsto f(\phi_y(x, t), y)$ is absolutely continuous on \mathbf{R}_+ for all (x, y) .

Averaging. Let us accelerate the dynamic of the component Y of the process by considering a transition rate matrix Q_ε proportional to $\frac{1}{\varepsilon}$, with $\varepsilon > 0$. In an equivalent way, one can as well consider that the Y -component evolves on a faster time-scale than the X -component by setting, for $t \in [0, T]$,

$$Y_\varepsilon(t) := Y\left(\frac{t}{\varepsilon}\right).$$

This accelerated process is denoted by $(X_\varepsilon(t), Y_\varepsilon(t))_{t \in [0, T]}$. The PDMP $(X_\varepsilon, Y_\varepsilon)$ constructed as described above has now two distinct time-scales. The variable Y_ε jumps on a fast time-scale of order ε between the states of \mathbf{Y} according to the X_ε -dependent transition matrix $Q_\varepsilon(X_\varepsilon)$ and between the jumps, the variable X_ε evolves on a slow time-scale of order 1 according to a system of ordinary differential equations

$$\frac{d}{dt}X_\varepsilon(t) = F_{Y_\varepsilon(t)}(X_\varepsilon(t)), \quad t \in [0, T].$$

The averaging problem consists in the study of the behaviour of the process when ε goes to zero, that is when the dynamic of the discrete component is infinitely accelerated. For this purpose, this discrete process needs to possess some kind of asymptotic stability as considered in the following assumption, see [5].

Assumption B. *For any $x \in \mathbf{R}^d$, the time-homogeneous Markov chain on \mathbf{Y} with generator $Q(x)$ is ergodic, that is, it visits with positive probability any state in \mathbf{Y} , for any starting point. We call $\pi(x)$ its unique invariant probability measure on \mathbf{Y} .*

Assumption B implies that, with x held fixed, the time homogeneous Markov chain with generator $Q(x)$ admits a unique invariant measure $\pi(x)$ to which the Markov chain converges as time goes to infinity. Moreover, $\pi_y(x) > 0$ for each $y \in \mathbf{Y}$. We call $\pi(x)$ the *quasistationary* measure associated to the component Y (when X is held fixed to x). As proved in [5], the quasistationary measure inherits the analytical properties of the vector fields and intensity rate functions. For instance, here, the quasistationary measure $\pi(x)$ is continuously differentiable in x under Assumption A.

The acceleration of the dynamic of the Y -component will induce an averaging of the component of the process with respect to the quasistationary measure π . We introduce the averaged vector field $\bar{F} : \mathbf{R}^d \rightarrow \mathbf{R}^d$ defined as

$$\bar{F}(x) := \sum_{y \in \mathbf{Y}} \pi_y(x) F_y(x). \quad (2)$$

Note that this averaged vector field is also continuously differentiable and bounded, thanks to Assumption A.1. In this context, the following averaging result holds.

Theorem 2. [5, Theorem 2.2] *Suppose that Assumptions A and B hold. Given $(x_0, y_0) \in \mathbf{R}^d \times \mathbf{Y}$, we denote by \bar{X} the unique solution of the Cauchy problem*

$$\frac{d}{dt}\bar{X}(t) = \bar{F}(\bar{X}(t)), \quad t \in [0, T]$$

with initial condition $\bar{X}(0) = x_0$. Then, for any $\delta > 0$, $y \in \mathbf{Y}$ and for any continuous function $f : [0, T] \rightarrow \mathbf{R}$ we have

$$\lim_{\varepsilon \rightarrow 0} \mathbf{P} \left(\left| \int_0^T f(t) \mathbf{1}_y(Y_\varepsilon(t)) dt - \int_0^T f(t) \pi_y(\bar{X}(t)) dt \right| \geq \delta \right) = 0,$$

and

$$\lim_{\varepsilon \rightarrow 0} \mathbf{P} \left(\sup_{t \in [0, T]} \|X_\varepsilon(t) - \bar{X}(t)\| \geq \delta \right) = 0.$$

This theorem states that the occupation measure associated to the fast component Y_ε converges in probability towards a measure having for density the \bar{X} -dependent quasistationary probability associated to Y_ε . In the meantime, there is also uniform convergence within the considered time window of the slow component X_ε , in probability, towards its averaged version \bar{X} . In [5], the author also obtained a large deviation principle associated to the considered process. We will not need this refinement in the present article.

2.2 Densities and averaging

We begin with a quite intuitive result about the existence of densities for the class of considered PDMP. We define for this purpose a measure γ on $\mathbf{R}^d \times \mathbf{Y}$ by

$$\gamma(\mathrm{d}x, \mathrm{d}y) := \lambda_d(\mathrm{d}x) \sum_{i=1}^{|\mathbf{Y}|} \delta_{y_i}(\mathrm{d}y).$$

We will require the initial value of the process to have a density with respect to γ .

Theorem 3. [21, Corollary 5.4] *Suppose that Assumptions A holds and that the initial value $(X(0), Y(0))$ of the PDMP has a density f_0 with respect to the measure γ . This density is assumed to be continuous in x . Then, for any $t \in [0, T]$, the couple $(X(t), Y(t))$ has a density f_t with respect to the measure γ , continuous in both x and t .*

We will provide some details on this theorem. Before that, for convenience, we state the hypothesis about the existence of a density with respect to ε .

Assumption C. *The initial value $(X_\varepsilon(0), Y_\varepsilon(0))$ of the PDMP possesses a density $f_{0,\varepsilon}$ with respect to the measure γ . This density is assumed to be continuous in x and with supremum norm uniformly bounded in ε .*

In general, if they exist, the densities $(f_t)_{t \in [0, T]}$ of a PDMP satisfy a transport equation as stated in Equations (2.26) and (2.27) in [9]. Indeed, according to Theorem 1, the Dynkin formula implies that for all bounded and measurable function g on $\mathbf{R}^d \times \mathbf{Y}$, let say continuously differentiable in x , we have,

$$\begin{aligned} & \int_{\mathbf{R}^d \times \mathbf{Y}} g(x, y) f_t(x, y) \gamma(\mathrm{d}x, \mathrm{d}y) - \int_{\mathbf{R}^d \times \mathbf{Y}} g(x, y) f_0(x, y) \gamma(\mathrm{d}x, \mathrm{d}y) \\ &= \int_0^t \int_{\mathbf{R}^d \times \mathbf{Y}} \nabla_x g(x, y) \cdot F_y(x) f_s(x, y) \gamma(\mathrm{d}x, \mathrm{d}y) \mathrm{d}s \\ & \quad + \int_0^t \int_{\mathbf{R}^d \times \mathbf{Y}} \sum_{z \in \mathbf{Y}} q_{yz}(x) [g(x, z) - g(x, y)] f_s(x, y) \gamma(\mathrm{d}x, \mathrm{d}y) \mathrm{d}s. \end{aligned}$$

This is a weak formulation for the following transport equation in conservative form,

$$\partial_t f_t(x, y) + \operatorname{div}(F_y(x) f_t(x, y)) = \sum_{z \in \mathbf{Y} \setminus \{y\}} q_{zy}(x) f_t(x, z), \quad t \in [0, T], \quad (x, y) \in \mathbf{R}^d \times \mathbf{Y}. \quad (3)$$

Assumption A as well as the continuity of the initial condition implies that the Cauchy problem associated to (3) has a unique solution continuous in space and time, see for example [4, Chapter III] or [19, Proposition 6.2]. From now on, we denote by $f_{t,\varepsilon}(x)$ this solution in the two-time scale setting, as soon as Assumptions A and C hold. Depending on the shape of the vector fields and intensity rate functions, the transport equation (3) is more or less difficult to solve explicitly. For our purpose, we do not need such explicit formula but will rather use some properties of the solution. Let us give a simple example.

Example 1. Assume that $F_y(x)$ and $q_{yz}(x)$ do not depend on x . Then, if Y is independent from $X(0)$,

$$f_t(x, y) = \mathbf{E} \left(f_0 \left(x - \int_0^t F_{Y(s)} \mathrm{d}s \right) \mathbf{1}_y(Y(t)) \right),$$

where f_0 denotes the density of $X(0)$ with respect to the Lebesgue measure on \mathbf{R}^d . We see that if f_0 is continuous and bounded in x , then so is $f_t(\cdot, y)$ for any $(t, y) \in [0, T] \times \mathbf{Y}$.

This example emphasizes that f_t transports the properties of f_0 with respect to x over time, as expected for the solution of a well posed transport equation. In the context of averaging, the behaviour of this regularity with respect to the parameter ε is of particular interest. To give the following result, we need to define the density associated to the average process \bar{X} . This density at time t , denoted by \bar{f}_t , satisfies the following transport equation in conservative form:

$$\begin{cases} \partial_t \bar{f}_t(x) + \operatorname{div}(\bar{f}_t(x)\bar{F}(x)) = 0, & x \in \mathbf{R}^d, t \in [0, T], \\ \bar{f}_0(x) = u_0(x), & x \in \mathbf{R}^d, \end{cases} \quad (4)$$

where u_0 is the density of $\bar{X}(0)$, assumed to be continuous. Then, since $\operatorname{div}\bar{F}$ is continuously differentiable on \mathbf{R}^d thanks to Assumption **A.1**, the Cauchy problem (4) has a unique solution which is continuous in time and space, see for example [19, Theorem 6.3].

Proposition 1. [16, Proposition 2.1] *Suppose that Assumptions A, B and C hold and moreover that for $\varepsilon \in (0, 1]$, the density $f_{0,\varepsilon}$ does not depend on ε , that is $f_{0,\varepsilon}(x, y) := f_0(x, y)$. Then, we have*

$$\lim_{\varepsilon \rightarrow 0} f_{t,\varepsilon}(x, y) = \bar{f}_t(x)\pi_y(x),$$

for any $(x, y, t) \in \mathbf{R}^d \times \mathbf{Y} \times [0, T]$, uniformly for x in any compact of \mathbf{R}^d , where $\bar{f}_t(x)$ is defined through the Cauchy problem (4) with $u_0 = f_0$.

Remark 1. The above result implies that for any $t \in [0, T]$ and any compact \mathcal{K} of \mathbf{R}^d , there exists ε_0 such that we have the following uniform bound,

$$\sup_{\varepsilon \in (0, \varepsilon_0]} \sup_{x \in \mathcal{K}} |f_{t,\varepsilon}(x)| < \infty.$$

As emphasized in the following example, continuing Example 1, Proposition 1 is quite intuitive when the process X_ε and Y_ε are fully decoupled.

Example 2. Assume that $F_y(x)$ and $q_{yz}(x)$ do not depend on x as in Example 1. Then, if Y_ε is independent from $X_\varepsilon(0)$ (whose density f_0 do not depend on ε),

$$f_{t,\varepsilon}(x, y) = \mathbf{E} \left(f_0 \left(x - \int_0^t F_{Y_\varepsilon(s)} ds \right) \mathbf{1}_y(Y_\varepsilon(t)) \right).$$

If f_0 is continuous, then the dominated convergence and ergodic theorems give,

$$\lim_{\varepsilon \rightarrow 0} f_{t,\varepsilon}(x, y) = f_0 \left(x - t \sum_{y \in \mathbf{Y}} \pi_y F_y \right) \pi_y$$

as required.

3 Averaging interaction with a crossing hypersurface and application to PDMP with boundary

3.1 A \mathcal{C}^1 -hypersurface and its crossings

We will consider the interplay of the process $(X_\varepsilon, Y_\varepsilon)$ with \mathbf{B} , a compact and connected \mathcal{C}^1 -hypersurface in \mathbf{R}^d . As such, this hypersurface borders a connected and bounded domain which we denote \mathbf{D} . The \mathcal{C}^1 -defining function associated to \mathbf{D} is denoted by ρ , see [14, Section 1.2]. We write $n_{\mathbf{B}}(x) := \frac{\nabla \rho(x)}{\|\nabla \rho(x)\|}$ for the outward unit normal of \mathbf{B} at $x \in \mathbf{B}$. Moreover, for any $\delta > 0$, we define the δ -tube around \mathbf{B} by

$$\mathcal{T}_\delta(\mathbf{B}) := \{x \in \mathbf{R}^d; d(x, \mathbf{B}) \leq \delta\}.$$

The Lebesgue surface measure associated to \mathbf{B} is denoted by $\sigma_{\mathbf{B}}$.

Assumption D. *The hypersurface \mathbf{B} may satisfy the following assumptions.*

D.1 *We assume that for any $x \in \mathbf{B}$ and $y \in \mathbf{Y}$,*

$$F_y(x) \cdot n_{\mathbf{B}}(x) \neq 0.$$

This implies that the trajectory of the flow are not tangents to the hypersurface \mathbf{B} .

D.2 *We assume that for any $x \in \mathbf{B}$,*

$$\bar{F}(x) \cdot n_{\mathbf{B}}(x) > 0.$$

In other words, under this assumption, if it starts from a point in \mathbf{D} , the hypersurface \mathbf{B} is fully attractive for the averaged process \bar{X} .

D.3 *We assume that for any $x \in \mathbf{B}$ and $y \in \mathbf{Y}$,*

$$F_y(x) \cdot n_{\mathbf{B}}(x) > 0.$$

In other words, under this assumption, if the euclidean component starts in \mathbf{D} , the hypersurface \mathbf{B} is fully attractive for X whatever the value of the discrete component Y .

Remark 2. Since \mathbf{B} is compact and the vector fields and outward normals are continuous functions, Assumptions D.2 and D.3 imply that there exists $\delta > 0$ such that

$$\inf_{(x,y) \in \mathcal{T}_{\delta}(\mathbf{B}) \times \mathbf{Y}} F_y(x) \cdot n_{\mathbf{B}}(x) > 0,$$

and

$$\inf_{x \in \mathcal{T}_{\delta}(\mathbf{B})} \bar{F}(x) \cdot n_{\mathbf{B}}(x) > 0.$$

Of course, Assumption D.3 implies Assumption D.2 (that implies Assumption D.1).

We proceed with the definition of the number of crossings of a hypersurface by a continuous process.

Definition 1 (Number of crossings). *Let Z be a process in $\mathcal{C}([0, T], \mathbf{R}^d)$. We say that Z crosses the hypersurface \mathbf{B} at time $s \in (0, T)$ if $Z(s) \in \mathbf{B}$ and there is $\varepsilon > 0$ such that $Z(u) \notin \mathbf{B}$ for $u \in (s - \varepsilon, s + \varepsilon) \setminus \{s\}$. For $t \in [0, T]$, we denote by $N_{\mathbf{B}}([0, t], Z)$ the number of crossings of the hypersurface \mathbf{B} by the process Z within the time window $[0, t]$. If the process Z is associated to another process Y valued in \mathbf{Y} , we denote by $N_{\mathbf{B}, y}([0, t], Z)$ the number of crossings of \mathbf{B} by Z when $Y(s) = y$ for any crossing time s .*

Remark 3. Using the defining function ρ , under Assumption D.1, a crossing of \mathbf{B} for the euclidean variable X is a crossing of 0 for the real valued process $\rho \circ X$.

The average number of crossings can be computed thanks to the Rice's formula as stated in the following proposition.

Proposition 2 (Rice's formula and averaging). *Suppose that Assumptions A, B, C and D.1 hold. Then, we have, for each $y \in \mathbf{Y}$,*

$$\mathbf{E}(N_{\mathbf{B}, y}([0, t], (X_{\varepsilon}, Y_{\varepsilon}))) = \int_{\mathbf{B}} |F_y(x) \cdot n_{\mathbf{B}}(x)| \int_0^t f_{s, \varepsilon}(x, y) ds \sigma_{\mathbf{B}}(dx). \quad (5)$$

Moreover,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}(N_{\mathbf{B}, y}([0, t], (X_{\varepsilon}, Y_{\varepsilon}))) = \int_{\mathbf{B}} |F_y(x) \cdot n_{\mathbf{B}}(x)| \pi_y \int_0^t \bar{f}_s(x) ds \sigma_{\mathbf{B}}(dx). \quad (6)$$

Proof. We already noticed that under the stated Assumption, the process has a density at any time, continuous in both space and time and uniformly bounded in ε , see Theorem 3, Proposition 1 and Remark 1. The process X_ε being continuous and piecewise continuously differentiable, the densities being continuous in both space and time, we can apply [2, Theorem 2.3] or [1, Corollary 2.11] to obtain the formula (5). Then, using Proposition 1 and the dominated convergence theorem, we obtain (6). \square

Remark 4. The above result does not imply that

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}(N_{\mathbf{B}}([0, t], X_\varepsilon)) = \mathbf{E}(N_{\mathbf{B}}([0, t], \bar{X}))$$

even if the family $\{X_\varepsilon\}_{\varepsilon \in (0,1]}$ is a family of continuous processes. Indeed, consider for example the case when $d = 1$. In this case the hypersurface reduces to a level $\mathbf{B} = \{b\}$. The measure $\sigma_{\mathbf{B}}$ is simply the counting measure and the normal to the surface can be taken to 1. Continuing Example 1, we have in this case, for $t \in [0, T]$ and $y \in \mathbf{Y}$,

$$\mathbf{E}(N_{\mathbf{B},y}([0, t], (X_\varepsilon, Y_\varepsilon))) = |F_y| \int_0^t \mathbf{E} \left(f_0 \left(b - \int_0^s F_{Y_\varepsilon(u)} du \right) \mathbf{1}_y(Y_\varepsilon(t)) \right) ds.$$

This converges to

$$|F_y| \pi_y \int_0^t f_0 \left(b - s \sum_{y \in \mathbf{Y}} \pi_y F_y \right) ds,$$

meaning that the probability that the euclidean process crosses the level b when the mode is equal to y is proportional to $\pi_y |F_y|$. Consider now that $\mathbf{Y} = \{-1, 1\}$, $\pi_{-1} = \pi_1 = 1/2$, $F_y = y$ and $b = 0$. Then

$$\mathbf{E}(N_{\mathbf{B},1}([0, t], (X_\varepsilon, Y_\varepsilon))) = \int_0^t \mathbf{E} \left(f_0 \left(- \int_0^s Y_\varepsilon(u) du \right) \mathbf{1}_1(Y_\varepsilon(t)) \right) ds$$

and also,

$$\mathbf{E}(N_{\mathbf{B},-1}([0, t], (X_\varepsilon, Y_\varepsilon))) = \int_0^t \mathbf{E} \left(f_0 \left(- \int_0^s Y_\varepsilon(u) du \right) \mathbf{1}_{-1}(Y_\varepsilon(t)) \right) ds$$

such that

$$\mathbf{E}(N_{\mathbf{B}}([0, t], X_\varepsilon)) = \int_0^t \mathbf{E} \left(f_0 \left(- \int_0^s Y_\varepsilon(u) du \right) \right) ds.$$

This converges to $tf_0(0)$ which is not equal to $\mathbf{E}(N_{\mathbf{B}}([0, t], \bar{X})) = 0$ since $\bar{X}(t) = X_0$ for any $t \in [0, T]$. The problem here is that the times between two crossings, one up and one down, cannot be uniformly minimised in epsilon. Assumptions D.2 and D.3 will prevent this phenomenon from happening.

Remark 5. The quantity

$$\int_0^t f_{s,\varepsilon}(x, y) ds$$

that appears in the given Rice's formula can be interpreted as the expectation of the local time spend by the process around $(x, y) \in \mathbf{B} \times \mathbf{Y}$, within the time window $[0, t]$, see [1].

Remark 6. Assume that the law of $X_\varepsilon(0)$ is supported in \mathbf{D} . Then, for $\varepsilon \in (0, 1]$, we can consider the number of up-crossings, denoted by $N_{\mathbf{B},y}^+([0, t], (X_\varepsilon, Y_\varepsilon))$ and the number of down-crossings, denoted by $N_{\mathbf{B},y}^-([0, t], (X_\varepsilon, Y_\varepsilon))$, of the set $\mathbf{B} \times \{y\}$ by the process $(X_\varepsilon, Y_\varepsilon)$. Suppose that Assumptions A, B, C and D.1 hold. Then, we have, for each $y \in \mathbf{Y}$,

$$\mathbf{E}(N_{\mathbf{B},y}^+([0, t], (X_\varepsilon, Y_\varepsilon))) = \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \int_0^t f_{s,\varepsilon}(x, y) ds \sigma_{\mathbf{B}}(dx)$$

and

$$\mathbf{E}(N_{\mathbf{B},y}^-([0,t], (X_\varepsilon, Y_\varepsilon))) = \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^- \int_0^t f_{s,\varepsilon}(x,y) ds \sigma_{\mathbf{B}}(dx),$$

where $x^+ = \max(x, 0)$ and $x^- = -\min(x, 0)$ for any real x .

We will now precise what the results of Proposition 2 implies on the behaviour of the slow-fast process around the crossing hypersurface. Let us now denote, for $\varepsilon \in (0, 1]$, by $T_\varepsilon^{\mathbf{B}}$ the first time of crossing of \mathbf{B} by X_ε :

$$T_\varepsilon^{\mathbf{B}} := \inf\{t \geq 0; X_\varepsilon(t) \in \mathbf{B}\}$$

with the usual convention that $\inf \emptyset = +\infty$.

Proposition 3. *Suppose that Assumptions A, B, C and D.3 hold. Then, for each $y \in \mathbf{Y}$,*

$$\lim_{\varepsilon \rightarrow 0} \mathbf{P}(Y_\varepsilon(T_\varepsilon^{\mathbf{B}}) = y; T_\varepsilon^{\mathbf{B}} \in [0, T]) = \int_{\mathbf{B}} F_y(x) \cdot n_{\mathbf{B}}(x) \pi_y(x) \int_0^T \bar{f}_s(x) ds \sigma_{\mathbf{B}}(dx). \quad (7)$$

More generally, for any measurable and bounded function h in $\mathbf{R}^d \times \mathbf{Y}$,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}(h(X_\varepsilon(T_\varepsilon^{\mathbf{B}}), Y_\varepsilon(T_\varepsilon^{\mathbf{B}})) \mathbf{1}_{[0,T]}(T_\varepsilon^{\mathbf{B}})) = \int_{\mathbf{B}} \sum_{y \in \mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \pi_y(x) \int_0^T \bar{f}_s(x) ds \sigma_{\mathbf{B}}(dx). \quad (8)$$

Proof. The proof is postponed to Section 5.1 □

When considering Assumption D.2 instead of D.3, only the up-crossings of the hypersurface will play a role.

Proposition 4. *Suppose that Assumptions A, B, C and D.2 hold. Assume that the law of $X_\varepsilon(0)$ is supported in \mathbf{D} . Then, the results of Proposition 3 hold. More precisely, for each $y \in \mathbf{Y}$,*

$$\lim_{\varepsilon \rightarrow 0} \mathbf{P}(Y_\varepsilon(T_\varepsilon^{\mathbf{B}}) = y; T_\varepsilon^{\mathbf{B}} \in [0, T]) = \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \pi_y(x) \int_0^T \bar{f}_s(x) ds \sigma_{\mathbf{B}}(dx).$$

More generally, for any measurable and bounded function h in $\mathbf{R}^d \times \mathbf{Y}$,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}(h(X_\varepsilon(T_\varepsilon^{\mathbf{B}}), Y_\varepsilon(T_\varepsilon^{\mathbf{B}})) \mathbf{1}_{[0,T]}(T_\varepsilon^{\mathbf{B}})) = \int_{\mathbf{B}} \sum_{y \in \mathbf{Y}} h(x, y) (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \pi_y(x) \int_0^T \bar{f}_s(x) ds \sigma_{\mathbf{B}}(dx).$$

Proof. The proof is postponed to Section 5.2. □

3.2 Averaging result for a class of piecewise deterministic process with boundary

The model. We are going to construct a PDMP $(X, Y) = (X(t), Y(t))_{t \in [0, T]}$ as in Section 2.1 but in constraining the euclidean component X to stay in the domain \mathbf{D} by making it jump when reaching the boundary \mathbf{B} . The euclidean component becomes a \mathbf{D} -valued piecewise continuous and deterministic process with only forced jumps when hitting the boundary \mathbf{B} . The Y -component is as before a \mathbf{Y} -valued piecewise constant process with purely stochastic jump times.

For $(x, y) \in \mathbf{D} \times \mathbf{Y}$, we define the hitting time of the boundary by the flow $\phi_y(x, \cdot)$ as

$$t_{\mathbf{B}}(x, y) := \inf\{t > 0; \phi_y(x, t) \in \mathbf{B}\}. \quad (9)$$

For any $(x, y) \in \mathbf{D} \times \mathbf{Y}$ we also define the survivor function

$$S_{(x,y)}(t) := \mathbf{1}_{[0, t_{\mathbf{B}}(x,y)]}(t) e^{-\int_0^t q_y(\phi_y(x,s)) ds}, \quad t \geq 0. \quad (10)$$

This is the survivor function of a non-negative random variable. We proceed to the construction of the process by recursion, as in Section 2.1. The only difference is the formulation of the post-jump locations. Let $\tau_0 = \sigma_0 = 0$ and $(X(0), Y(0)) = (X_0, Y_0)$ be a $\mathbf{D} \times \mathbf{Y}$ -valued random variable. Assume that the process is built until time τ_{n-1} , for $n \in \mathbf{N}$. Then, we can define a random variable σ_n , the inter-jump time, satisfying

$$\mathbf{P}(\sigma_n > t | (X_{n-1}, Y_{n-1}) = (x, y)) = S_{(x,y)}(t), \quad t \geq 0.$$

We define the n th jump time by

$$\tau_n := \tau_{n-1} + \sigma_n$$

and we set

$$(X(t), Y(t)) = \begin{cases} (\phi_{Y_{n-1}}(X_{n-1}, t - \tau_{n-1}), Y_{n-1}) & \text{for } \tau_{n-1} \leq t < \tau_n, \\ (X_n, Y_n) & \text{for } t = \tau_n. \end{cases}$$

The n th post-jump location (X_n, Y_n) is a $\mathbf{D} \times \mathbf{Y}$ -valued random variable such that with $\tilde{X}_n = \phi_{Y_{n-1}}(X_{n-1}, \sigma_n)$, the distribution of X_n on \mathbf{D} is given by

$$\nu_y(dw | \tilde{X}_n) \mathbf{1}_{\mathbf{B}}(\tilde{X}_n) + \delta_{\tilde{X}_n}(dw) \mathbf{1}_{\mathbf{B}^c}(\tilde{X}_n)$$

where, for $(x, y) \in \mathbf{B} \times \mathbf{Y}$, $\nu_y(\cdot | x)$ is a probability measure on \mathbf{D} . The distribution of Y_n on \mathbf{Y} is given by

$$\delta_{Y_{n-1}}(dy) \mathbf{1}_{\mathbf{B}}(\tilde{X}_n) + \sum_{z \in \mathbf{Y} \setminus Y_{n-1}} \frac{q_{Y_{n-1}, z}(\tilde{X}_n)}{q_{Y_{n-1}}(\tilde{X}_n)} \delta_z(dy) \mathbf{1}_{\mathbf{B}^c}(\tilde{X}_n).$$

That is to say that either X or Y jumps at the jump times, but not both. As in Section 2.1, the component Y is a jump process that may jump at X -dependent rates. The component X evolves according to a family of vector fields parametrized by Y and has only forced jumps when reaching the boundary, if it does so. Indeed, let us define the times of these forced jumps. We set $T_0^{\mathbf{B}} := 0$ and for $i \in \mathbf{N}$,

$$T_i^{\mathbf{B}} := \inf\{t > T_{i-1}^{\mathbf{B}}; X(t^-) \in \mathbf{B}\}$$

with $X(t^-) := \lim_{s \nearrow t} X(s^-)$. By construction of the process, \mathbf{P} -a.s., we have

$$\frac{dX(t)}{dt} = F_{Y(t)}(X(t)), \quad t \in [T_{i-1}^{\mathbf{B}}, T_i^{\mathbf{B}}), \quad i \in \mathbf{N}.$$

The process X satisfies a switching ordinary differential equation with jumps at the boundary. A detailed example is provided in Section 4.

Definition 2. Let $t \in [0, T]$. The number of jumps of X until time t is

$$p_{\mathbf{B}}(t) = \sum_{i=1}^{\infty} \mathbf{1}_{[0, t]}(T_i^{\mathbf{B}}).$$

The counting process $p_{\mathbf{B}}$ counts the number of jumps from the boundary of the process X .

Proposition 5. Suppose that Assumptions A holds and that

$$\mathbf{E}(p_{\mathbf{B}}(T)) < \infty.$$

Then, the process (X, Y) is well defined.

Proof. See [3, Assumption 24.4 and Proposition 24.6]. □

Remark 7. The assumption

$$\mathbf{E}(p_{\mathbf{B}}(T)) < \infty$$

is satisfied if, for instance, the jump measure has support in \mathbf{D} far enough from \mathbf{B} . Indeed, in such a case, the process X can only reach the boundary a finite number of times over a finite time horizon, almost-surely. More precisely, assume that there exists $\delta > 0$ such that, for any $(x, y) \in \mathbf{B} \times \mathbf{Y}$, the jump measure at the boundary $\nu_y(\cdot | x)$ has support outside the semi-tube $\mathcal{T}_{\delta}(\mathbf{B}) \cap \mathbf{D}$. Then, in this case, the process can not reach more than $\lfloor \frac{T}{\delta} \sup_{(x,y) \in \mathbf{D} \times \mathbf{Y}} \|F_y(x)\| \rfloor$ of times the boundary \mathbf{B} , where $\lfloor \cdot \rfloor$ stands for the floor function.

Acceleration Replacing the intensity operator Q by $\frac{1}{\varepsilon}Q$ in the construction of the process, we obtain a two-time scale process $(X_\varepsilon(t), Y_\varepsilon(t))_{t \in [0, T]}$ as in Section 2.1. Its counting measure at the boundary will be denoted by $p_{\mathbf{B}, \varepsilon}$. Our aim is still to study its behaviour when ε goes to zero, that is when the dynamic of the discrete component is infinitely accelerated.

Definition 3 (Averaged jump measure at the boundary and averaged field). *Suppose that Assumptions B and D.2 hold. Recall that the averaged field is defined, for $x \in \bar{\mathbf{D}}$, by*

$$\bar{F}(x) := \sum_{y \in \mathbf{Y}} F_y(x) \pi_y(x).$$

For $(x, y) \in \mathbf{B} \times \mathbf{Y}$, we define

$$\mu_y^{\mathbf{B}}(x) := \frac{\pi_y(x) (F_y(x) \cdot n_{\mathbf{B}}(x))^+}{\sum_{z \in \mathbf{Y}} \pi_z(x) (F_z(x) \cdot n_{\mathbf{B}}(x))^+}$$

where π is defined in Assumption B. The averaged jump measure at the boundary is given by

$$\bar{\nu}(d\tilde{x}|x) := \sum_{y \in \mathbf{Y}} \nu_y(d\tilde{x}|x) \mu_y^{\mathbf{B}}(x).$$

Remark 8. Note that, thanks to Assumption D.2, this average field is positive around the boundary and that for each y , $\mu_y^{\mathbf{B}}$ is well defined.

The averaged jump measure at the boundary is thus a weighting of the boundary jump measures at the boundary, with the weights taking into account both the invariant measure associated with the process Y and the interactions between the flows and the boundary, as can be expected from the crossing formula given in Proposition 2.

Assumption E. *For any $(x, y) \in \mathbf{B} \times \mathbf{Y}$, the measure $\nu_y(\cdot|x)$ is absolutely continuous with respect to the Lebesgue measure on \mathbf{D} , with bounded and continuous density. Moreover, we assume that the supports of these densities are in $\{x \in \mathbf{D}; d(x, \mathbf{B}) \geq \delta\}$ for some $\delta > 0$.*

According to Remark 7, Assumption E implies that the number of jumps of this PDMP is finite almost-surely.

All the processes that we consider are càdlàg \mathbf{R}^d -valued processes defined on $[0, T]$, that is, belong to the Skorokhod space $\mathbf{D}([0, T], \mathbf{R}^d)$. We consider this space endowed with the so-called J_1 topology, see [11] and the references therein. With these notations and assumptions we obtain the following result.

Theorem 4. *Suppose that Assumptions A, B, C, D.2 and E hold. Assume moreover that $X_\varepsilon(0)$ converges in law towards some random variable \bar{X}_0 , when ε goes to zero. Then, the process $(X_\varepsilon(t))_{t \in [0, T]}$ converges in law, when ε goes to zero, towards the averaged process $(\bar{X}(t))_{t \in [0, T]}$ starting at \bar{X}_0 and such that:*

1. *The process \bar{X} is piecewise continuously differentiable with jumps at times $(\bar{T}_i^{\mathbf{B}})_{i \in \mathbf{N}}$ such that for $i \in \mathbf{N}$,*

$$\bar{T}_i^{\mathbf{B}} := \inf\{t > \bar{T}_{i-1}^{\mathbf{B}}; \bar{X}(t^-) \in \mathbf{B}\}$$

with the conventions that $\bar{T}_0^{\mathbf{B}} := 0$.

2. *In between two jumps, for $i \in \mathbf{N}$ and $t \in [\bar{T}_{i-1}^{\mathbf{B}}, \bar{T}_i^{\mathbf{B}})$ the process X satisfies the equation*

$$\frac{d}{dt} \bar{X}(t) = \bar{F}(\bar{X}(t)), \tag{11}$$

with initial condition $\bar{X}(\bar{T}_{i-1}^{\mathbf{B}})$.

3. *At jump times, for $i \in \mathbf{N}$, $\bar{X}(\bar{T}_i^{\mathbf{B}})$ is distributed according to the jump measure $\bar{\nu}(\cdot | \bar{X}(\bar{T}_i^{\mathbf{B}, -}))$.*

Proof. The proof is postponed to Section 5.3. □

4 Example and possible extensions

We start this section with an example showing to what kind of situation our main result can be applied, this is Section 4.1 . We then go on to discuss some possible extensions of our main result in Section 4.2.

4.1 A constrained centrifugal motion in dimension two

Consider the unit circle

$$\mathbf{B} := \{x \in \mathbf{R}^2; \|x\| = 1\}$$

with as corresponding domain the unit open disc

$$\mathbf{D} := \{x \in \mathbf{R}^2; \|x\| < 1\}$$

and outward normal $n_{\mathbf{B}}(x) = x$. For a finite set of angles \mathbf{Y} in $(-\frac{\pi}{2}, \frac{\pi}{2})$, let say

$$\mathbf{Y} := \left\{ -\frac{\pi}{2} + \frac{k}{n}\pi; k \in \{1, \dots, n-1\} \right\}$$

with $n \in \mathbf{N}$, $n \geq 3$, we define the rotation matrices

$$\forall y \in \mathbf{Y}, \quad R_y := \begin{pmatrix} \cos(y) & -\sin(y) \\ \sin(y) & \cos(y) \end{pmatrix}.$$

For $y \in \mathbf{Y}$, the vector field is given by

$$\forall x \in \overline{\mathbf{D}}, \quad F_y(x) := \lambda R_y x,$$

with $\lambda > 0$. These vector fields satisfy Assumption **A.1**. Due to the fact that $\mathbf{Y} \subset (-\frac{\pi}{2}, \frac{\pi}{2})$, the motion of the X -component will be centrifugal until it reaches the boundary. Indeed, for any $(x, y) \in \overline{\mathbf{D}} \times \mathbf{Y}$, $x \neq 0$,

$$F_y(x) \cdot n_{\mathbf{B}}(x) = \lambda \cos(y) \|x\|^2 > 0.$$

Assumption D.3 is thus satisfied.

The transition rate functions are defined, for $(y, z) \in \mathbf{Y}^2$, $y \neq z$ and $x \in \overline{\mathbf{D}}$ by

$$Q_{yz}(x) := \frac{1}{2}(1 + \|x\|^2)\mathbf{1}_{|y-z|=1}.$$

Theses transition rates functions satisfy Assumption **A.2**. When x is held fixed, the associated invariant probability $\pi(x)$ is the uniform probability on \mathbf{Y} : Assumption B is verified. For $x \in \mathbf{B}$ and $y \in \mathbf{Y}$, the jump measure $\nu_y(dw|x)$ at the boundary, has same law as

$$Ux + (1 - U)(x - 2 \cos(\Theta_y)R_{\Theta_y}x)$$

where:

- U is a Beta law supported by $[1/4, 3/4]$ and with parameter $(2, 2)$. Its density is denoted by f_U ;
- Θ_y is a Beta law supported by $[\min(0, y), \max(0, y)]$ if y is not equal to zero and $[-\pi/(2n); \pi/(2n)]$ when $y = 0$, always with parameter $(2, 2)$. Its density is denoted by f_{Θ_y} .

The random variables U and Θ_y are assumed to be independent. This jump measure means that when it hits the boundary, the process is rejected further into the centre of the domain, in a random direction between the direction of the normal to the contact point and the direction of the velocity vector at the contact point (which therefore depends on y). We have, by change of variables,

$$\nu_y(dw|x) = f_U(\psi_1(w|x))f_{\Theta_y}(\psi_2(w|x))|\det J_\psi(w|x)|dw$$

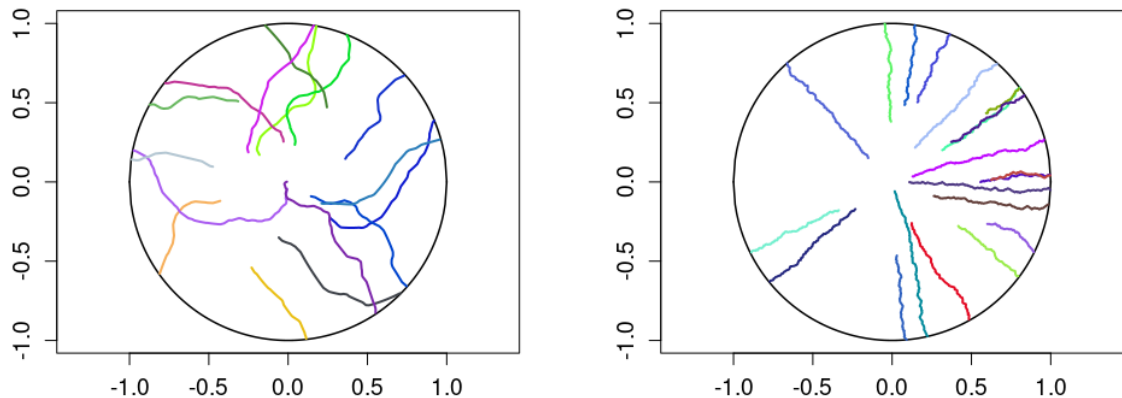


Figure 1: Left: trajectory of the process with $\lambda = 0.1$, $n = 10$ and $\varepsilon = 1$ up to the horizon time $T = 300$. We coloured each piece of the trajectory between two forced jumps with a different colour. Right: same thing but with $\varepsilon = 0.001$. For the simulation of piecewise deterministic Markov processes, we refer to [20].

where, for w such that $x \cdot w \neq 1$,

$$\begin{aligned} \psi(w|x) &:= (\psi_1(w|x), \psi_2(w|x)) \\ &:= \left(1 - \frac{1}{2} \frac{\|x - w\|^2}{1 - x \cdot w}, \tan^{-1} \left(\frac{x \cdot w^\perp}{1 - x \cdot w} \right) \right) \end{aligned}$$

with $w^\perp = \begin{pmatrix} -w_2 \\ w_1 \end{pmatrix}$. Remark that if w satisfies $x \cdot w = 1$, then w is not in the support of $\nu_y(\cdot|x)$. Indeed,

$$x \cdot w = 1 \Leftrightarrow x \cdot (w - x) = 0,$$

so that w is on the tangent line to \mathbf{B} at the point x . Assumption E is satisfied for this family of jump measures.

Therefore, if the initial value of the euclidean process has a density with respect to the Lebesgue measure on \mathbf{D} and that this one is continuous, we can apply Theorem 4. The average vector field is given, for $x \in \overline{\mathbf{D}}$, by

$$\bar{F}(x) = \left(\frac{\lambda}{n-1} \sum_{k=1}^{n-1} R_{-\frac{\pi}{2} + \frac{k}{n}\pi} \right) x = \frac{\lambda}{(n-1) \tan(\pi/(2n))} x$$

and the averaged measure at the boundary by

$$\bar{\nu}(dw|x) = \sum_{y \in \mathbf{Y}} f_{\Theta_y}(\psi_2(w|x)) \mu_y^{\mathbf{B}}(x) f_U(\psi_1(w|x)) |\det J_\psi(w|x)| dw,$$

with

$$\mu_y^{\mathbf{B}}(x) = \tan \left(\frac{\pi}{2n} \right) \cos(y),$$

which is here independent of x . Two trajectories of the process are displayed in Figure 1, for two different ε , showing the effect of acceleration.

4.2 Two possible extensions

In this section we present two natural extensions of Theorem 4. They require, to show them, only minor modifications of the convergence proof presented in Section 5.3.

4.2.1 Change of domain after a forced jump

In this variant, it is assumed that the X -component of the process can change domain after each forced jump. Let $(\mathbf{D}_i)_{i \geq 0}$ be a sequence of domains in \mathbf{R}^d with corresponding sequence of smooth boundaries $(\mathbf{B}_i)_{i \geq 0}$. The random variable X_0 is valued in \mathbf{D}_0 . Then, as before,

$$T_1^{\mathbf{B}_1} := \inf\{t > 0; X(t^-) \in \mathbf{B}_0\}$$

but the measure $\nu_{Y(T_1^{\mathbf{B}_1, -})}(\cdot | X(T_1^{\mathbf{B}_1, -}))$ is now valued in the domain \mathbf{D}_1 . This should be emphasized by writing $\nu_{1, Y(T_1^{\mathbf{B}_1, -})}(\cdot | X(T_1^{\mathbf{B}_1, -}))$. And so on and so forth. Our assumptions, in particular the Assumptions D.3, B, C and E must be adapted accordingly. For example, Assumption D.2 becomes as follows, taking into account the successive and possibly different boundaries.

Assumption 1 (Attractive boundary). *For any $i \geq 1$, we assume that*

$$\forall x \in \mathbf{B}_i, \quad \bar{F}(x) \cdot n_{\mathbf{B}_i}(x) > 0.$$

We could as well change the form of the flow after each forced jumps by indexing the vector fields as the successive domains. This change of domain framework is the one considered in [3] for the definition of a general PDMP.

4.2.2 Slow-fast discrete process

Another natural framework is the one where the discrete component exhibits itself a slow fast dynamic. In this case, the transition rate matrix is of the form

$$Q^{(S)} + \frac{1}{\varepsilon} Q^{(F)}$$

where S and F stand for Slow and Fast. We assume that there exists a partition $(\mathbf{Y}_j)_{1 \leq j \leq N}$, with $N \geq 2$, of the discrete space \mathbf{Y} such that the transitions in between the classes are given by the transition rate matrix Q_S and inside a specific class j by the bloc $Q_j^{(F)}$ of the transition rate matrix $Q^{(F)}$, see [22, Section 3.6].

This multi-scale framework fits naturally to some applications, as for conductance based neuron models in neuroscience [16]. This framework is fully developed in [22] in the case of continuous time Markov chains where other kind of applications are considered.

In contrary to the extension presented in the previous section, the Assumption B to be made in order to deal with this multi-scale case requires further reformulation. We follow [16].

Assumption 2. *We assume that for $x \in \mathbf{D}$ and $i \in \{1, \dots, N\}$, there exists a unique probability $\pi^{(i)}(x)$ on \mathbf{Y}_i solution of $\pi^{(i)}(x) Q_i^{(F)}(x) = 0$.*

In this multi-scale setting, the averaged process is still a switching ordinary differential equation (\bar{X}, \bar{Y}) with jumps at the boundary where the state belonging to fast transition clusters have been averaged. Namely, the process \bar{Y} is valued in $\{1, \dots, N\}$ and jumps at x -dependent rates $\bar{Q}_{ij}(x)$ such that for $i \neq j$,

$$\bar{Q}_{ij}(x) = \sum_{(y,z) \in \mathbf{Y}_i \times \mathbf{Y}_j} Q_{yz}^{(S)}(x) \pi_y^{(i)}(x).$$

The averaged euclidean component \bar{X} evolves according to a family of averaged vector fields given, for $x \in \mathbf{D}$ and $i \in \{1, \dots, N\}$, by

$$\bar{F}_i(x) = \sum_{y \in \mathbf{Y}_i} F_y(x) \pi_y^{(i)}(x).$$

Assumption D.2 becomes, for $i \in \{1, \dots, N\}$,

$$\forall x \in \mathbf{B}, \quad \bar{F}_i(x) \cdot n_{\mathbf{B}}(x) > 0.$$

Then, when \bar{X} hits the boundary, it jumps according to the averaged jump measure at the boundary given, for $x \in \mathbf{B}$ and $i \in \{1, \dots, N\}$, by

$$\bar{\nu}_i(\cdot|x) = \frac{\sum_{y \in \mathbf{Y}_i} \nu_y(\cdot|x) \pi_y^{(i)}(x) (F_y(x) \cdot n_{\mathbf{B}}(x))^+}{\sum_{y \in \mathbf{Y}_i} \pi_y^{(i)}(x) (F_y(x) \cdot n_{\mathbf{B}}(x))^+}.$$

5 Proofs

This section is devoted to the proofs of the main results of the present article.

5.1 Proof of Proposition 3

For any $\varepsilon > 0$ and each $y \in \mathbf{Y}$, we have,

$$\mathbf{P}(Y_\varepsilon(T_\varepsilon^{\mathbf{B}}) = y; T_\varepsilon^{\mathbf{B}} \in [0, T]) = \mathbf{P}(N_{\mathbf{B},y}([0, T], (X_\varepsilon, Y_\varepsilon)) = 1),$$

since under Assumption D.3, the process X_ε may cross the hypersurface at most one time. We then remark that

$$\mathbf{P}(N_{\mathbf{B},y}([0, T], (X_\varepsilon, Y_\varepsilon)) = 1) = \mathbf{E}(N_{\mathbf{B},y}([0, T], (X_\varepsilon, Y_\varepsilon))),$$

and apply Proposition 2 to get Equation (7). To get the formula (8) in the proposition, we need to show that for any measurable and bounded function h and for any $\varepsilon > 0$, we have,

$$\mathbf{E}(h(X_\varepsilon(T_\varepsilon^{\mathbf{B}}), Y_\varepsilon(T_\varepsilon^{\mathbf{B}})) \mathbf{1}_{[0, T]}(T_\varepsilon^{\mathbf{B}})) = \int_{\mathbf{B}} \sum_{y \in \mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \int_0^T f_{\varepsilon, s}(x, y) ds \sigma_{\mathbf{B}}(dx).$$

For this purpose, remark that under the stated assumptions, see [2],

$$\begin{aligned} & \mathbf{E}(h(X_\varepsilon(T_\varepsilon^{\mathbf{B}}), Y_\varepsilon(T_\varepsilon^{\mathbf{B}})) \mathbf{1}_{[0, T]}(T_\varepsilon^{\mathbf{B}})) \\ &= \mathbf{E} \left(\lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_0^T h(X_\varepsilon(s), Y_\varepsilon(s)) F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot n_{\mathbf{B}}(X_\varepsilon(s)) \|\nabla \rho(X_\varepsilon(s))\| \mathbf{1}_{[-\delta, \delta]}(\rho(X_\varepsilon(s))) ds \right). \end{aligned}$$

Since we have the following uniform in δ domination, (recall that we have at most one crossing),

$$\begin{aligned} & \frac{1}{2\delta} \int_0^T h(X_\varepsilon(s), Y_\varepsilon(s)) F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot n_{\mathbf{B}}(X_\varepsilon(s)) \|\nabla \rho(X_\varepsilon(s))\| \mathbf{1}_{[-\delta, \delta]}(\rho(X_\varepsilon(s))) ds \\ & \leq \sup_{(x, y) \in \mathbf{R}^d \times \mathbf{Y}} |h(x, y)|, \end{aligned}$$

the dominated convergence theorem yields,

$$\begin{aligned} & \mathbf{E}(h(X_\varepsilon(T_\varepsilon^{\mathbf{B}}), Y_\varepsilon(T_\varepsilon^{\mathbf{B}})) \mathbf{1}_{[0, T]}(T_\varepsilon^{\mathbf{B}})) \\ &= \lim_{\delta \rightarrow 0} \mathbf{E} \left(\frac{1}{2\delta} \int_0^T h(X_\varepsilon(s), Y_\varepsilon(s)) F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot n_{\mathbf{B}}(X_\varepsilon(s)) \|\nabla \rho(X_\varepsilon(s))\| \mathbf{1}_{[-\delta, \delta]}(\rho(X_\varepsilon(s))) ds \right). \end{aligned}$$

Using the densities to express the expectation we obtain, using also the Fubini theorem,

$$\begin{aligned} & \mathbf{E}(h(X_\varepsilon(T_\varepsilon^{\mathbf{B}}), Y_\varepsilon(T_\varepsilon^{\mathbf{B}}))\mathbf{1}_{[0,T]}(T_\varepsilon^{\mathbf{B}})) \\ &= \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{\mathbf{R}^d} \sum_{y \in \mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \int_0^T f_{s,\varepsilon}(x, y) ds \|\nabla \rho(x)\| \mathbf{1}_{[-\delta, \delta]}(\rho(x)) dx. \end{aligned}$$

Then, since \mathbf{B} is a bounded \mathcal{C}^1 -hypersurface and ρ its \mathcal{C}^1 -defining function,

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{\mathbf{R}^d} \sum_{y \in \mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \int_0^T f_{s,\varepsilon}(x, y) ds \|\nabla \rho(x)\| \mathbf{1}_{[-\delta, \delta]}(\rho(x)) dx \\ &= \int_{\mathbf{B}} \sum_{y \in \mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \int_0^T f_{s,\varepsilon}(x, y) ds \sigma_{\mathbf{B}}(dx). \end{aligned}$$

The result follows using the dominated convergence theorem and Proposition 1.

5.2 Proof of Proposition 4

For the proof of Proposition 3 to be still valid replacing Assumption D.3 by Assumption D.2, the number of crossings of the hypersurface by X_ε must be uniformly bounded in ε , at least for small ε . This is the case thanks to Assumption D.2. To see this, we proceed by contradiction. Let us assume that there exists a sequence (ε_k) going to zero such that for any k large enough we have $N_{\mathbf{B}}([0, T], X_{\varepsilon_k}) \geq 3$. By Proposition 7 and the Skorokhod's representation theorem, one can assume that (X_{ε_k}) converges to \bar{X} uniformly on $[0, T]$, almost-surely. The fact that for any k large enough X_{ε_k} crosses \mathbf{B} at least three times implies that \bar{X} crosses \mathbf{B} and it does it exactly one time by Assumption D.2. Remark that this also implies, always for k large enough, that the number of up-crossings of \mathbf{B} by X_{ε_k} is equal to the number of down crossings of \mathbf{B} plus one. In between the first down-crossing $T_{\text{down}, \varepsilon_k}$ of \mathbf{B} and the last up-crossing $T_{\text{last}, \varepsilon_k}$, the increasing and decreasing phases of $\rho(X_{\varepsilon_k}(\cdot))$ must compensate. That is to say, for k large enough,

$$\begin{aligned} & \mathbf{E} \left(\sum_{y \in \mathbf{Y}} \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \int_{T_{\text{down}, \varepsilon_k}}^{T_{\text{last}, \varepsilon_k}} f_{s, \varepsilon_k}(x, y) ds \sigma_{\mathbf{B}}(dx) \right) \\ &= \mathbf{E} \left(\sum_{y \in \mathbf{Y}} \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^- \int_{T_{\text{down}, \varepsilon_k}}^{T_{\text{last}, \varepsilon_k}} f_{s, \varepsilon_k}(x, y) ds \sigma_{\mathbf{B}}(dx) \right). \end{aligned}$$

We then have, for k large enough,

$$\mathbf{E} \left(\sum_{y \in \mathbf{Y}} \int_{\mathbf{B}} F_y(x) \cdot n_{\mathbf{B}}(x) \int_{T_{\text{down}, \varepsilon_k}}^{T_{\text{last}, \varepsilon_k}} f_{s, \varepsilon_k}(x, y) ds \sigma_{\mathbf{B}}(dx) \right) = 0$$

and therefore, when k goes to infinity, by dominated convergence,

$$\mathbf{E} \left(\sum_{y \in \mathbf{Y}} \int_{\mathbf{B}} F_y(x) \cdot n_{\mathbf{B}}(x) \pi_y(x) \bar{f}_{\bar{T}^{\mathbf{B}}}(x) \sigma_{\mathbf{B}}(dx) \right) = 0,$$

where $\bar{T}^{\mathbf{B}}$ is the crossing time of \mathbf{B} by \bar{X} . The last equality is

$$\mathbf{E} \left(\int_{\mathbf{B}} \bar{F}(x) \cdot n_{\mathbf{B}}(x) \bar{f}_{\bar{T}^{\mathbf{B}}}(x) \sigma_{\mathbf{B}}(dx) \right) = 0. \quad (12)$$

Under Assumption D.2, we have

$$\overline{F}(x) \cdot n_{\mathbf{B}}(x) > 0, \quad \forall x \in \mathbf{B}.$$

Moreover, since \overline{X} crosses \mathbf{B} , we also have $f_{\overline{F}\mathbf{B}}(x) > 0$ in some neighbourhood of \mathbf{B} . These two facts are in contradiction with Equation (12).

For small enough ε , there is thus at most one crossing of \mathbf{B} by X_ε and we can apply the same reasoning as in the proof of Proposition 4.

5.3 Proof of Theorem 4

We follow the program initiated by Prokhorov, see [11]: we prove that the family $\{X_\varepsilon, \varepsilon \in (0, 1]\}$ is tight in $\mathbf{D}([0, T], \mathbf{R}^d)$ in Section 5.3.1 and then identify the limit in Section 5.3.2.

5.3.1 Tightness

The proofs of tightness for slow-fast processes with boundary under general conditions could be complicated because of the presence of fast stochastic jumps combined with the presence of instantaneous jumps at the boundary, that is the presence of forced jumps. One standard technique used to overcome this difficulty is the penalization method as it is exposed in [15, 7]. Here, Assumption C allows for the use of a more direct and simple approach: the existence of a density for the process at any time allows to control the probability that this one is in any set by the size of this precise set.

We use the tightness criterion presented in [13, Theorem 16.11]. For $\varepsilon > 0$, since the process X_ε evolves in \mathbf{D} which is bounded, there exists a constant C such that

$$\sup_{\varepsilon \in (0, 1]} \sup_{t \in [0, T]} \|X_\varepsilon(t)\| \leq C, \quad \mathbf{P}\text{-a.s.}$$

Let δ and M be two positive reals. For any $\varepsilon > 0$, let $t \in [0, T]$ and $h > 0$ be such that $t + h \leq T$. Thanks to Assumptions C and E, there is at most one forced jump between times t and $t + h$ for h small enough. Thus, we write,

$$\|X_\varepsilon(t + h) - X_\varepsilon(t)\| \leq h \sup_{(x, y) \in \mathbf{D} \times \mathbf{Y}} \|F_y(x)\| + \text{diam}(\mathbf{D}) \mathbf{1}_{\exists s \in (t, t+h], X_\varepsilon(s^-) \in \mathbf{B}}.$$

Under Assumptions C and E, [9, Theorem 2.5 and Section 5.1] implies that for all $\varepsilon > 0$, the random variable $X_\varepsilon(t)$ has a density with respect to the Lebesgue measure for any $t \in [0, T]$ which is uniformly bounded in ε . Therefore,

$$\mathbf{P}(\exists s \in (t, t + h], X_\varepsilon(s^-) \in \mathbf{B}) \leq \mathbf{P}(X_\varepsilon(t) \in \mathcal{T}_{\sup_{(x, y) \in \mathbf{D} \times \mathbf{Y}} \|F_y(x)\| h}(\mathbf{B})) = \mathcal{O}(h),$$

uniformly in $\varepsilon \in (0, 1]$. The family $\{X_\varepsilon; \varepsilon \in (0, 1]\}$ is thus tight in $\mathbf{D}([0, T], \mathbf{R}^d)$, endowed with the Skorokhod topology.

5.3.2 Identification of the limit.

Let \overline{X} be an accumulation point of the tight family $\{X_\varepsilon; \varepsilon \in (0, 1]\}$. Our aim is to show that, for any $h \in \mathbf{R}^d$, we have

$$\begin{aligned} \mathbf{E}(\overline{X}(t) \cdot h) &= \mathbf{E}(\overline{X}(0) \cdot h) + \mathbf{E} \left(\int_0^t \overline{F}(\overline{X}(s)) \cdot h ds \right) \\ &\quad + \mathbf{E} \left(\int_0^t \int_{\mathbf{D}} [w - \overline{X}(s^-)] \cdot h \overline{\nu}(dw | \overline{X}(s^-)) \overline{p}^{\mathbf{B}}(ds) \right), \end{aligned}$$

for any $t \in [0, T]$, as this will characterize the law of \bar{X} , as stated in [3, Theorem 26.14]. We will still denote by $(X_\varepsilon)_\varepsilon$ a sequence converging towards \bar{X} in the Skorokhod topology. By convergence in the Skorokhod topology, we have, for any point of continuity $t \in [0, T]$ of \bar{X} :

$$\mathbf{E}(\bar{X}(t) \cdot h) = \lim_{\varepsilon \rightarrow 0} \mathbf{E}(X_\varepsilon(t) \cdot h).$$

Following [3, Theorem 26.14], we also know that for any $\varepsilon \in (0, 1]$ and $t \in [0, T]$,

$$\begin{aligned} \mathbf{E}(X_\varepsilon(t) \cdot h) &= \mathbf{E}(X_\varepsilon(0) \cdot h) + \mathbf{E} \left(\int_0^t F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot h ds \right) \\ &\quad + \mathbf{E} \left(\int_0^t \int_{\mathbf{D}} [w - X_\varepsilon(s^-)] \cdot h \nu_{Y_\varepsilon(s^-)}(dw | X_\varepsilon(s^-)) \bar{p}_\varepsilon^{\mathbf{B}}(ds) \right). \end{aligned}$$

We begin with the limit of the two first terms in the right hand side of the above expression, as they do not present any difficulty.

Proposition 6. *We have,*

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}(X_\varepsilon(0) \cdot h) = \mathbf{E}(\bar{X}(0) \cdot h)$$

and for any $t \in [0, T]$,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E} \left(\int_0^t F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot h ds \right) = \mathbf{E} \left(\int_0^t \bar{F}(\bar{X}(s)) \cdot h ds \right).$$

Proof. The first equality holds by assumption: there is convergence in law at time zero. For the second one, since the counting measure of the forced jumps does not appear in the considered expression, this is a direct consequence of the convergence in the Skorokhod topology of (X_ε) towards \bar{X} and of standard averaging techniques for piecewise deterministic Markov processes without forced jumps, as presented in [16]. \square

It remains to show that

$$\begin{aligned} &\lim_{\varepsilon \rightarrow 0} \mathbf{E} \left(\int_0^t \int_{\mathbf{D}} [w - X_\varepsilon(s^-)] \cdot h \nu_{Y_\varepsilon(s^-)}(dw | X_\varepsilon(s^-)) \bar{p}_\varepsilon^{\mathbf{B}}(ds) \right) \\ &= \mathbf{E} \left(\int_0^t \int_{\mathbf{D}} [w - \bar{X}(s^-)] \cdot h \bar{\nu}(dw | \bar{X}(s^-)) \bar{p}^{\mathbf{B}}(ds) \right) \end{aligned}$$

Since we can not expect the convergence of the stopping times in the Skorokhod topology, we need to take a different route to show this convergence. To proceed, observe that we have, for $\varepsilon \in (0, 1]$ and $t \in [0, T]$,

$$\mathbf{E} \left(\int_0^t \int_{\mathbf{D}} [w - X_\varepsilon(s^-)] \cdot h \nu_{Y_\varepsilon(s^-)}(dw | X_\varepsilon(s^-)) \bar{p}_\varepsilon^{\mathbf{B}}(ds) \right) = \sum_{i=1}^{\infty} \mathbf{E} \left(\Gamma(X_\varepsilon(T_{i,\varepsilon}^{\mathbf{B},-}), Y_\varepsilon(T_{i,\varepsilon}^{\mathbf{B},-})) \mathbf{1}_{T_{i,\varepsilon}^{\mathbf{B}} \leq t} \right),$$

where, for $(x, y) \in \mathbf{B} \times \mathbf{Y}$,

$$\Gamma(x, y) := \int_{\mathbf{D}} [w - x] \cdot h \nu_y(dw | x).$$

Since Assumption E holds, note that the sum over i is finite, \mathbf{P} -a.s. We are going to show that for any $i \in \mathbf{N}$,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E} \left(\Gamma(X_\varepsilon(T_{i,\varepsilon}^{\mathbf{B},-}), Y_\varepsilon(T_{i,\varepsilon}^{\mathbf{B},-})) \mathbf{1}_{T_{i,\varepsilon}^{\mathbf{B}} \leq t} \right) = \mathbf{E} \left(\bar{\Gamma}(\bar{X}(T_i^{\mathbf{B},-})) \mathbf{1}_{\bar{T}_i^{\mathbf{B}} \leq t} \right),$$

with, for $x \in \mathbf{B}$,

$$\bar{\Gamma}(x) := \sum_{y \in \mathbf{Y}} \Gamma(x, y) \mu_y^{\mathbf{B}}(x)$$

where $\mu^{\mathbf{B}}$ is defined in Definition 3. The proof will repose on two main ingredients:

1. The first one is the consideration of companion processes that follow the law of $(X_\varepsilon, Y_\varepsilon)$ but stop to jump after the i th forced jump.
2. These companion processes allow us to make the link with the second main ingredient: the Rice's formulas to count the average number of continuous crossings of a hypersurface for a piecewise smooth process, as presented in Proposition 3.

We move on to the definition of these companion processes.

Definition 4 (Companion processes). *For $i \in \mathbf{N}$ and $\varepsilon \in (0, 1]$, we denote by $(X_\varepsilon^{(i)}, Y_\varepsilon^{(i)})$ the process constructed as $(X_\varepsilon, Y_\varepsilon)$ until the i th jumps of $X_\varepsilon^{(i)}$. Then, if $T_{i,\varepsilon}^{(i),\mathbf{B}}$ is finite, for $t \geq T_{i,\varepsilon}^{(i),\mathbf{B}}$,*

$$\frac{d}{dt} X_\varepsilon^{(i)}(t) = F_{Y_\varepsilon^{(i)}(s)}(X_\varepsilon^{(i)}(s)) \mathbf{1}_{\mathbf{D} \cup \mathcal{T}_\delta(\mathbf{B})}(X_\varepsilon^{(i)}(s)),$$

with $\delta > 0$ chosen as in Remark 2.

In particular, there is no forced jump after the i th forced jumps, the process $X^{(i)}$ remaining continuous after this time. Usual averaging results allow to consider the limit of the first companion process.

Proposition 7 (Averaging principle when there is no forced jumps). *Under the Assumptions A and B, the family of continuous processes $(X_\varepsilon^{(1)})_{\varepsilon \in (0,1]}$ converges in law towards a process $\bar{X}^{(1)}$ on $\mathcal{C}([0, T], \mathbf{R}^d)$ endowed with the uniform topology. The process $\bar{X}^{(1)}$ is such that for $t \in [0, T]$,*

$$\bar{X}^{(1)}(t) = \bar{X}_0 + \int_0^t \bar{F}(\bar{X}^{(1)}(s)) ds$$

where \bar{F} is given in Definition 3.

Proof. Under the assumptions made, the result is given by Proposition 7. □

Remark 9. When $i \geq 2$, the same result holds true for the processes $(X_\varepsilon^{(i)})_{\varepsilon \in (0,1]}$ when far from the boundary. Indeed, for $i \geq 2$, we can show that the family $(X_\varepsilon^{(i)})_{\varepsilon \in (0,1]}$ is tight in $\mathbf{D}([0, T], \mathbf{R}^d)$ quite exactly as for $(X_\varepsilon)_{\varepsilon \in (0,1]}$ and that for almost every $t \in [0, T]$, in fact for all t except at the jump times,

$$\frac{d}{dt} \bar{X}^{(i)}(t) = \bar{F}(\bar{X}^{(i)}(t)).$$

We are in position to show that for any $i \in \mathbf{N}$ and $t \in [0, T]$,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E} \left(\Gamma(X_\varepsilon(T_{i,\varepsilon}^{\mathbf{B},-}), Y_\varepsilon(T_{i,\varepsilon}^{\mathbf{B},-})) \mathbf{1}_{T_{i,\varepsilon}^{\mathbf{B},-} \leq t} \right) = \mathbf{E} \left(\bar{\Gamma}(\bar{X}(\bar{T}_i^{\mathbf{B},-})) \mathbf{1}_{\bar{T}_i^{\mathbf{B},-} \leq t} \right),$$

We begin with the first term where $i = 1$.

Proposition 8. *We have*

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E} \left(\Gamma(X_\varepsilon(T_{1,\varepsilon}^{\mathbf{B},-}), Y_\varepsilon(T_{1,\varepsilon}^{\mathbf{B},-})) \mathbf{1}_{T_{1,\varepsilon}^{\mathbf{B},-} \leq t} \right) = \mathbf{E} \left(\bar{\Gamma}(\bar{X}(\bar{T}_1^{\mathbf{B},-})) \mathbf{1}_{\bar{T}_1^{\mathbf{B},-} \leq t} \right)$$

and for any $t \in [0, T]$,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{P}(T_{1,\varepsilon}^{\mathbf{B},-} \leq t) = \mathbf{P}(\bar{T}_1^{\mathbf{B},-} \leq t).$$

Proof. Let us remark that the probability that X^ε has its first jump with Y_ε equals to y is equal to the probability that the companion process $X_\varepsilon^{(1)}$ crosses \mathbf{B} for the first time with $Y_\varepsilon^{(1)}$ equals to y . The result is then a direct consequence of Proposition 4. □

Proposition 9. For $i = 2$

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E} \left(\Gamma(X_\varepsilon(T_{2,\varepsilon}^{\mathbf{B},-}), Y_\varepsilon(T_{2,\varepsilon}^{\mathbf{B},-})) \mathbf{1}_{T_{2,\varepsilon}^{\mathbf{B}} \leq t} \right) = \mathbf{E} \left(\bar{\Gamma}(\bar{X}(\bar{T}_2^{\mathbf{B},-})) \mathbf{1}_{\bar{T}_2^{\mathbf{B}} \leq t} \right).$$

Proof. We have, following the same idea as in the proof of the Proposition 8, for $\varepsilon \in (0, 1]$ and $t \in [0, T]$,

$$\mathbf{P}(Y_\varepsilon(T_{2,\varepsilon}^{\mathbf{B},-}) = y; T_{2,\varepsilon}^{\mathbf{B}} \leq t) = \mathbf{E}(N_{\mathbf{B},y}([T_{1,\varepsilon}^{\mathbf{B},(2),-}, t], X_\varepsilon^{(2)}) \mathbf{1}_{T_{1,\varepsilon}^{\mathbf{B},(2)} \leq t}),$$

where $T_{1,\varepsilon}^{\mathbf{B},(2)}$ is the time of the first forced jump for the second companion process. It has same law as $T_{1,\varepsilon}^{\mathbf{B}}$ or $T_{1,\varepsilon}^{\mathbf{B},(1)}$, the first crossing time of the boundary for the first companion process. Conditioning on the last forced jump time, we observe that

$$\begin{aligned} \mathbf{E}(N_{\mathbf{B},y}([T_{1,\varepsilon}^{\mathbf{B},(2),-}, t], X_\varepsilon^{(2)})) &= \mathbf{E} \left(\mathbf{E} \left(N_{\mathbf{B},y} \left([T_{1,\varepsilon}^{\mathbf{B},(2),-}, t], X_\varepsilon^{(2)} \right) \middle| \mathcal{F}_{T_{1,\varepsilon}^{\mathbf{B},(2),-}} \right) \mathbf{1}_{T_{1,\varepsilon}^{\mathbf{B},(2)} \leq t} \right) \\ &= \int_0^t \mathbf{E}(N_{\mathbf{B},y}([s, t], X_\varepsilon^{(2)})) \frac{d}{ds} \mathbf{P}(T_{1,\varepsilon}^{\mathbf{B},(2)} \leq s) ds \\ &= \int_0^t \mathbf{E}(N_{\mathbf{B},y}([s, t], X_\varepsilon^{(2)})) \frac{d}{ds} \mathbf{P}(T_{1,\varepsilon}^{\mathbf{B},(1)} \leq s) ds, \end{aligned}$$

where we have used the fact that the two first companion processes have the same law until the first forced jump of the second companion process. In the above formula, we have

$$\frac{d}{ds} \mathbf{P}(T_{1,\varepsilon}^{\mathbf{B},(1)} \leq s) = \sum_{y \in \mathbf{Y}} \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^+ f_{s,\varepsilon}(x, y) \sigma_{\mathbf{B}}(dx).$$

At the limit when ε goes to 0, we obtain, by dominated convergence,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E}(N_{\mathbf{B},y}([T_{1,\varepsilon}^{\mathbf{B},-}, t], X_\varepsilon^{(2)})) = \int_0^t \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \pi_y(x) \int_s^t f_{\bar{X}(u)}(x) du \sigma^{\mathbf{B}}(dx) \frac{d}{ds} \mathbf{P}(\bar{T}_1^{\mathbf{B}} \leq s) ds.$$

We can then proceed as in the proof of Proposition 3 to generalize this formula to Γ :

$$\begin{aligned} &\lim_{\varepsilon \rightarrow 0} \mathbf{E} \left(\Gamma(X_\varepsilon(T_{2,\varepsilon}^{\mathbf{B},-}), Y_\varepsilon(T_{2,\varepsilon}^{\mathbf{B},-})) \mathbf{1}_{T_{2,\varepsilon}^{\mathbf{B}} \leq t} \right) \\ &= \int_0^t \int_{\mathbf{B}} \sum_{y \in \mathbf{Y}} \Gamma(x, y) (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \pi_y(x) \int_s^t f_{\bar{X}(u)}(x) du \sigma^{\mathbf{B}}(dx) \frac{d}{ds} \mathbf{P}(\bar{T}_1^{\mathbf{B}} \leq s) ds. \end{aligned}$$

This last expression is precisely $\mathbf{E} \left(\bar{\Gamma}(\bar{X}(\bar{T}_2^{\mathbf{B},-})) \mathbf{1}_{\bar{T}_2^{\mathbf{B}} \leq t} \right)$, as required. \square

Reproducing the reasoning of the proof of Proposition 9, we obtain by recurrence that for any $i \in \mathbf{N}$,

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E} \left(\Gamma(X_\varepsilon(T_{i,\varepsilon}^{\mathbf{B},-}), Y_\varepsilon(T_{i,\varepsilon}^{\mathbf{B},-})) \mathbf{1}_{T_{i,\varepsilon}^{\mathbf{B}} \leq t} \right) = \mathbf{E} \left(\bar{\Gamma}(\bar{X}(\bar{T}_i^{\mathbf{B},-})) \mathbf{1}_{\bar{T}_i^{\mathbf{B}} \leq t} \right).$$

The proof of Theorem 4 is thus complete.

References

- [1] AZAÏS, Romain et GENADOT, Alexandre. Estimation of the average number of continuous crossings for non-stationary non-diffusion processes. *Journal of Statistical Planning and Inference*, 2019, vol. 198, p. 119-138.
- [2] BOROVKOV, Konstantin et LAST, Günter. On Rice's formula for stationary multivariate piecewise smooth processes. *Journal of Applied Probability*, 2012, vol. 49, no 2, p. 351-363.
- [3] DAVIS, Mark HA. *Markov models and optimization*. Routledge, 2018.
- [4] ENGEL, Klaus-Jochen et NAGEL, Rainer. One-parameter semigroups for linear evolution equations. In : *Semi-group forum*. Springer-Verlag, 2001. p. 278-280.
- [5] FAGGIONATO, Alessandra, GABRIELLI, Davide, et CRIVELLARI, M. Ribezzi. Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors. *Markov Processes and Related Fields*, 2010, vol. 16, 497-548.
- [6] GENADOT, Alexandre. A multiscale study of stochastic spatially-extended conductance-based models for excitable systems. 2013. Thèse de doctorat. Université Pierre et Marie Curie-Paris VI.
- [7] GENADOT, Alexandre. Averaging for some simple constrained Markov processes. *Probability and Mathematical Statistics*, 2019, vol. 39, no 1, p. 139-158.
- [8] GERSTNER, Wulfram, KISTLER, Werner M., NAUD, Richard, et al. *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press, 2014.
- [9] GWIŹDŹ, Piotr et TYRAN-KAMIŃSKA, Marta. Densities for piecewise deterministic Markov processes with boundary. *Journal of Mathematical Analysis and Applications*, 2019, vol. 479, no 1, p. 384-425.
- [10] HAIRER, Martin et LI, Xue-Mei. Averaging dynamics driven by fractional Brownian motion. *The Annals of Probability*, 2020, vol. 48, no 4, p. 1826-1860.
- [11] JAKUBOWSKI, Adam. The Skorokhod space in functional convergence: a short introduction. In : *International conference: Skorokhod Space*. 2007. p. 11-18.
- [12] KAC, Mark. On the average number of real roots of a random algebraic equation. *Bulletin of the American Mathematical Society*, 1943, vol. 49, no 4, p. 314-320.
- [13] KALLENBERG, Olav. *Foundations of modern probability*. New York : springer, 1997.
- [14] KRANTZ, Steven G. et PARKS, Harold R. *The geometry of domains in space*. Springer Science & Business Media, 1999.
- [15] KURTZ, Thomas G. Martingale problems for constrained Markov processes. *Recent advances in stochastic calculus*, 1990, p. 151-168.
- [16] PAKDAMAN, Khashayar, THIEULLEN, Michèle, et WAINRIB, Gilles. Asymptotic expansion and central limit theorem for multiscale piecewise-deterministic Markov processes. *Stochastic Processes and their Applications*, 2012, vol. 122, no 6, p. 2292-2318.
- [17] PARDOUX, E. et VERETENNIKOV, A. Yu. On the Poisson equation and diffusion approximation. I. *Annals of probability*, 2001, p. 1061-1085.
- [18] PAVLIOTIS, Grigoris et STUART, Andrew. *Multiscale methods: averaging and homogenization*. Springer Science & Business Media, 2008.
- [19] PERTHAME, Benoît. *Transport equations in biology*. Springer Science & Business Media, 2006.
- [20] RIEDLER, Martin G. Almost sure convergence of numerical approximations for piecewise deterministic Markov processes. *Journal of Computational and Applied Mathematics*, 2013, vol. 239, p. 50-71.
- [21] TYRAN-KAMIŃSKA, Marta. Substochastic semigroups and densities of piecewise deterministic Markov processes. *Journal of mathematical analysis and applications*, 2009, vol. 357, no 2, p. 385-402.
- [22] YIN, George G. et ZHANG, Qing. *Continuous-time Markov chains and applications: a singular perturbation approach*. Springer, 2012.