AVERAGING FOR SLOW–FAST PIECEWISE DETERMINISTIC MARKOV PROCESSES WITH AN ATTRACTIVE BOUNDARY

ALEXANDRE GÉNADOT,* *Univ. Bordeaux, CNRS, INRIA, Bordeaux INP, IMB, UMR 5251*

Abstract

In this paper we consider the problem of averaging for a class of piecewise deterministic Markov processes (PDMPs) whose dynamic is constrained by the presence of a boundary. On 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 PDMPs.

Keywords: Constrained Markov process; multiscale modelling; singular perturbation; number of crossings

2020 Mathematics Subject Classification: Primary 60F99
Secondary 60H99; 60J99

1. Introduction

In this article we consider piecewise deterministic Markov processes (henceforth PDMPs), introduced by Davis [7] in their Euclidean-mode setting. In this framework, a PDMP consists of an ordered pair $(X(t), Y(t))_{t\geq 0}$ on $\mathbb{R}^d \times \mathbf{Y}$, 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 $\mathbb{R}^d$, 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 it reaches the hypersurface $\mathbf{B}$ considered as a boundary. We are thus in the framework of averaging for constrained Markov processes. A typical trajectory of the Euclidean variable is displayed in Figure 1.

Averaging for unconstrained Markov process, that is, without the presence of a boundary, has been studied by several authors for decades and is well understood for a wide variety of Markov processes. For instance, see [29] for continuous-time Markov chains, [23] for diffusion processes, [15] for fractional Brownian motion, [9], [10], and [22] for piecewise deterministic Markov processes, and [24] 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
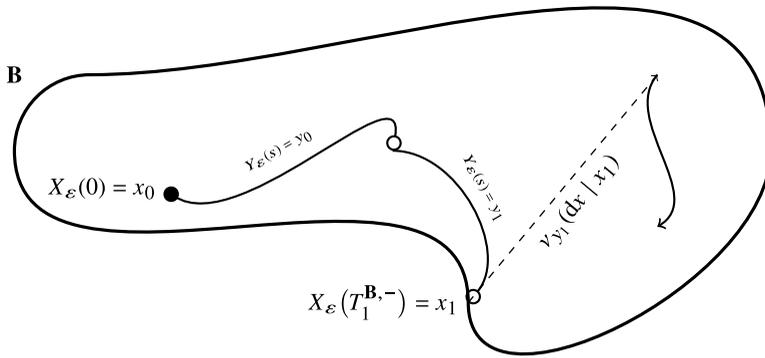
FIGURE 1. A typical trajectory of the Euclidean variable X_ε. The process starts at time 0 from $X_\varepsilon(0) = x_0$ and follows some flow parametrized by the initial mode $Y_\varepsilon(0) = y_0$. At some random time, Y_ε jumps to a new value y_1, causing the flow of the Euclidean variable to change. On hitting the boundary $\mathbf{B}$ at time $T_1^{\mathbf{B}}$ and at some point x_1, the Euclidean variable jumps inside the domain delimited by $\mathbf{B}$ according to the mode-parametrized jump measure $\nu_{y_1}(dx \mid x_1)$. Then it continues its course according to the current flow. And so on and so forth.

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 [22] for applications of this method in neuroscience.

To the best of our knowledge, averaging for constrained Markov processes, i.e. with the presence of a boundary, has not been the object of so much investigation, particularly 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 method', described in [20, Section 6.4]. This method consists in considering a penalized process jumping at a 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 the authors of [5] set up a general method to study Markov processes with constraints. They called this method the 'patchwork martingale problem'. It has been successfully applied to reflected diffusions [6].

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. For the 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 of the average number of crossings for piecewise smooth processes performed in [1] and [3]. In these articles, Rice's formula for the average number of crossings of a hypersurface is obtained, and the so-called crossing measure at the boundary is studied more specifically in [3]. One of the main assumptions in these articles is the existence of densities for the process at any time. The problem of the existence of densities for PDMPs has been extensively studied in [14] and [28]. In our setting, a set of sufficient assumptions

consists in supposing 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 is described, we are able to handle the presence of forced jumps in the dynamic of the PDMP, that is, to tackle the problem of averaging for a constrained PDMP.

Our main result is stated in Theorem 4, Section 3.2. We obtain the convergence in law of the constrained two time-scale PDMP $(X_\varepsilon, Y_\varepsilon)$ towards an averaged process that is still a constrained PDMP. However, this PDMP has only forced jumps at the boundary. The corresponding averaged vector field and averaged jump measure governing its dynamic are fully described. The averaged vector field corresponds, as in unconstrained averaging, to the vector field governing the dynamic of the Euclidean variable X averaged against the X-dependent invariant probability measure $\pi(X)$ associated to the mode process Y. The form of the averaged jump measure at the boundary takes into account the interplay between the flow, the boundary, and the jump measure. Indeed, the averaged jump measure corresponds to an averaging of the mode-dependent jump measures (ν_y in Figure 1) against the measure $\pi(X)$ weighted by the strength of the collision of the Euclidean variable against the boundary **B**. In the expression of these weights, this appears through the presence of the scalar product between the outward normal vector of **B** and the vector field associated to the Euclidean variable.

The present article is organized as follows. Section 2 is devoted to the study of PDMPs without boundary. In Section 2.1 we present 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 make 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 along with two possible extensions of our results and an application to a conductance-based neuron model. Section 5.3 is devoted to the proofs, and in particular to the proof of our averaging theorem for PDMPs with an attractive boundary.

2. Averaging for a class of PDMPs without boundary

2.1. The model and its averaging

Notations and general assumptions. As usual, $\mathbb{R}$ is the set of real numbers, with $\mathbb{R}_+^*$ the set of positive real numbers. The space $\mathbb{R}^d$, where $d \geq 2$ is an integer, is endowed with the Euclidean distance d with corresponding norm $\| \cdot \|$ and associated scalar product between points x and $\tilde{x}$ in $\mathbb{R}^d$ denoted by $x \cdot \tilde{x}$. $\mathbb{N}$ is the set of positive integers. We let λ_d denote the Lebesgue measure on $\mathbb{R}^d$. In the following, $\mathbf{Y} = \{y_1, \ldots, y_{|\mathbf{Y}|}\}$ 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 constructed 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 $\mathbb{R}^d$. Moreover, we assume that there exist constants $\kappa_1, \kappa_2 > 0$ such that*

$$\|F_y(x)\| \leq \kappa_1 + \kappa_2 \|x\| \quad \textit{for all } x \in \mathbb{R}^d.$$

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

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

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 $q_{yz} : x \in \mathbb{R}^d \mapsto q_{yz}(x) \in \mathbb{R}_+^*$ 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 $x \in \mathbb{R}^d$, 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 $\mathbb{R}^d \times \mathbf{Y}$ as follows. The continuous component, X, takes values in $\mathbb{R}^d$ 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 $(x, y) \in \mathbb{R}^d \times \mathbf{Y}$ we define the survivor function

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

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 $\mathbb{R}^d \times \mathbf{Y}$-valued random variable. Assume that the process is built until time τ_{n-1}, for $n \in \mathbb{N}$. Then we can define a random variable σ_n, the inter-jump time, satisfying

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

We define the nth 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 nth post-jump location (X_n, Y_n) is a $\mathbb{R}^d \times \mathbf{Y}$-valued random variable such that

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

and

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

We have thus constructed 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. (Theorem 26.14 of [7].) *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* $f : \mathbb{R}^d \times \mathbf{Y} \to \mathbb{R}$*, such that* $t \mapsto f(\phi_y(x, t), y)$ *is absolutely continuous on* $\mathbb{R}_+$ *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 $1/\varepsilon$, with $\varepsilon > 0$. In an equivalent way, one can also 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 now has 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 the system of ordinary differential equations

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

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 [9].

Assumption B. *For any* $x \in \mathbb{R}^d$*, 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 let* $\pi(x)$ *denote 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 [9], 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 $\overline{F} : \mathbb{R}^d \to \mathbb{R}^d$ defined as

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

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. (Theorem 2.2 of [9].) *Suppose that Assumptions* A *and* B *hold. Given* $(x_0, y_0) \in \mathbb{R}^d \times \mathbf{Y}$, *we let* $\overline{X}$ *denote the unique solution of the Cauchy problem*

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

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

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

and

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

This theorem states that the occupation measure associated to the fast component Y_ε converges in probability towards a measure having for density the $\overline{X}$-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 $\overline{X}$. Faggionato *et al.* [9] 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 fairly intuitive result about the existence of densities for the class of PDMPs considered. For this purpose we define a measure γ on $\mathbb{R}^d \times \mathbf{Y}$ 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. (Corollary 5.4 of [28].) *Suppose that Assumption* 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 pair* $(X(t), Y(t))$ *has a density* f_t *with respect to the measure* γ, *continuous in both x and t.*

We will provide some details of 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 [14]. Indeed, according to Theorem 1, the Dynkin formula implies that for all bounded and measurable function g on $\mathbb{R}^d \times \mathbf{Y}$, let us say continuously

differentiable in x, we have

$$
\int_{\mathbb{R}^d \times \mathbf{Y}} g(x, y) f_t(x, y) \gamma(\mathrm{d}x, \mathrm{d}y) - \int_{\mathbb{R}^d \times \mathbf{Y}} g(x, y) f_0(x, y) \gamma(\mathrm{d}x, \mathrm{d}y)
$$

$$
= \int_0^t \int_{\mathbb{R}^d \times \mathbf{Y}} \nabla_x g(x, y) \cdot F_y(x) f_s(x, y) \gamma(\mathrm{d}x, \mathrm{d}y)\, \mathrm{d}s
$$

$$
+ \int_0^t \int_{\mathbb{R}^d \times \mathbf{Y}} \sum_{z \in \mathbf{Y}} q_{yz}(x)[g(x, z) - g(x, y)] f_s(x, y) \gamma(\mathrm{d}x, \mathrm{d}y)\, \mathrm{d}s.
$$

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

$$
\partial_t f_t(x, y) + \mathrm{div}(F_y(x) f_t(x, y)) = \sum_{z \in \mathbf{Y} \setminus \{y\}} q_{zy}(x) f_t(x, z), \quad t \in [0, T], \ (x, y) \in \mathbb{R}^d \times \mathbf{Y}. \quad (1)
$$

Assumption A as well as the continuity of the initial condition implies that the Cauchy problem associated to (1) has a unique solution continuous in space and time; see for example [8, Chapter III] or [25, Proposition 6.2]. From now on, we let $f_{t,\varepsilon}(x)$ denote 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 (1) is more or less difficult to solve explicitly. For our purpose, we do not need such an explicit formula but will instead 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 of $X(0)$,

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

where f_0 denotes the density of $X(0)$ with respect to the Lebesgue measure on $\mathbb{R}^d$. 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 $\overline{X}$. This density at time t, denoted by $\overline{f}_t$, satisfies the following transport equation in conservative form:

$$
\begin{cases}
\partial_t \overline{f}_t(x) + \mathrm{div}(\overline{f}_t(x)\overline{F}(x)) = 0, & x \in \mathbb{R}^d, \ t \in [0, T], \\
\overline{f}_0(x) = u_0(x), & x \in \mathbb{R}^d,
\end{cases} \quad (2)
$$

where u_0 is the density of $\overline{X}(0)$, assumed to be continuous. Then, since $\mathrm{div}\overline{F}$ is continuously differentiable on $\mathbb{R}^d$ thanks to Assumption A.1, the Cauchy problem (2) has a unique solution which is continuous in time and space; see for example [25, Theorem 6.3].

Proposition 1. (Proposition 2.1 of [22].) *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*

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

for any $(x, y, t) \in \mathbb{R}^d \times \mathbf{Y} \times [0, T]$, *uniformly for* x *in any compact of* $\mathbb{R}^d$, *where* $\bar{f}_t(x)$ *is defined via the Cauchy problem* (2) *with* $u_0 = f_0$.

Remark 1. The above result implies that for any $t \in [0, T]$ and any compact $\mathcal{K}$ of $\mathbb{R}^d$, 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 fairly intuitive when the processes 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 of $X_\varepsilon(0)$ (whose density f_0 does not depend on ε),

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

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

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

as required.

3. Averaging interaction with a crossing hypersurface and application to PDMPs 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 $\mathbb{R}^d$. As such, this hypersurface borders a connected and bounded domain which we denote by $\mathbf{D}$. The $\mathcal{C}^1$-defining function associated to $\mathbf{D}$ is denoted by ρ; see [19, 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

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

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 trajectories of the flow are not tangents to the hypersurface $\mathbf{B}$.

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

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

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

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 **D**, the hypersurface **B** is fully attractive for X whatever the value of the discrete component Y.*

Remark 2. Since **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

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

Of course, Assumption D.3 implies Assumption D.2 (which 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 $\mathcal{C}([0, T], \mathbb{R}^d)$. We say that Z crosses the hypersurface **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 let $N_{\mathbf{B}}([0, t], Z)$ denote the number of crossings of the hypersurface **B** by the process Z within the time window [0, t]. If the process Z is associated to another process Y valued in **Y**, we let $N_{\mathbf{B},y}([0, t], Z)$ denote the number of crossings of **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 **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 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 $y \in \mathbf{Y}$,*

$$\mathbf{E}(N_{\mathbf{B},y}([0, t], (X_\varepsilon, Y_\varepsilon))) = \int_{\mathbf{B}} |F_y(x) \cdot n_{\mathbf{B}}(x)| \int_0^t f_{s,\varepsilon}(x, y) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x). \tag{3}$$

Moreover,

$$\lim_{\varepsilon \to 0} \mathbf{E}(N_{\mathbf{B},y}([0, t], (X_\varepsilon, Y_\varepsilon))) = \int_{\mathbf{B}} |F_y(x) \cdot n_{\mathbf{B}}(x)| \pi_y \int_0^t \overline{f}_s(x) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x). \tag{4}$$

Proof. We have 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 is continuous and piecewise continuously differentiable, and the densities are continuous in both space and time, so we can apply [3, Theorem 2.3] or [1, Corollary 2.11] to obtain the formula (3). Then, using Proposition 1 and the dominated convergence theorem, we obtain (4).   □

Remark 4. The above result does not imply that

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

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 be 1. Continuing Example 1, we have in this case, for $t \in [0, T]$ and $y \in \mathbf{Y}$,

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

This converges to

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

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

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

and also

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

such that

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

This converges to $t f_0(0)$, which is not equal to $\mathbf{E}(N_{\mathbf{B}}([0, t], \overline{X})) = 0$ since $\overline{X}(t) = X_0$ for any $t \in [0, T]$. The problem here is that the times between two crossings, one up and one down, cannot be uniformly minimized in epsilon. Assumptions D.2 and D.3 will prevent this phenomenon from happening.

Remark 5. The quantity

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

that appears in Rice's formula can be interpreted as the expectation of the local time spent 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 **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 $y \in \mathbf{Y}$,

$$\mathbf{E}\big(N^+_{\mathbf{B},y}([0, t], (X_\varepsilon, Y_\varepsilon))\big) = \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \int_0^t f_{s,\varepsilon}(x, y) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x)$$

and

$$\mathbf{E}\big(N^-_{\mathbf{B},y}([0, t], (X_\varepsilon, Y_\varepsilon))\big) = \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^- \int_0^t f_{s,\varepsilon}(x, y) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x),$$

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

We will now explain what the results of Proposition 2 imply about the behaviour of the slow–fast process around the crossing hypersurface. Let $T^{\mathbf{B}}_\varepsilon$, for $\varepsilon \in (0, 1]$, denote the first time of crossing of **B** by X_ε:

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

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

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

$$\lim_{\varepsilon \to 0} \mathbf{P}\big(Y_\varepsilon(T^{\mathbf{B}}_\varepsilon) = y; T^{\mathbf{B}}_\varepsilon \in [0, T]\big) = \int_{\mathbf{B}} F_y(x) \cdot n_{\mathbf{B}}(x) \pi_y(x) \int_0^T \overline{f}_s(x) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x). \tag{5}$$

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

$$\lim_{\varepsilon \to 0} \mathbf{E}\big(h\big(X_\varepsilon(T^{\mathbf{B}}_\varepsilon), Y_\varepsilon(T^{\mathbf{B}}_\varepsilon)\big)\mathbf{1}_{[0,T]}\big(T^{\mathbf{B}}_\varepsilon\big)\big)$$

$$= \int_{\mathbf{B}} \sum_{y \in \mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \pi_y(x) \int_0^T \overline{f}_s(x) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x). \tag{6}$$

Proof. The proof is postponed to Section 5.1. $\qquad\square$

When considering Assumption D.1 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 **D**. Then the results of Proposition 3 hold. More precisely, for each $y \in \mathbf{Y}$,*

$$\lim_{\varepsilon \to 0} \mathbf{P}\big(Y_\varepsilon(T^{\mathbf{B}}_\varepsilon) = y; T^{\mathbf{B}}_\varepsilon \in [0, T]\big) = \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \pi_y(x) \int_0^T \overline{f}_s(x) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x).$$

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

$$\lim_{\varepsilon \to 0} \mathbf{E}\big(h\big(X_\varepsilon(T^{\mathbf{B}}_\varepsilon), Y_\varepsilon(T^{\mathbf{B}}_\varepsilon)\big)\mathbf{1}_{[0,T]}\big(T^{\mathbf{B}}_\varepsilon\big)\big)$$

$$= \int_{\mathbf{B}} \sum_{y \in \mathbf{Y}} h(x, y)(F_y(x) \cdot n_{\mathbf{B}}(x))^+ \pi_y(x) \int_0^T \overline{f}_s(x) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x).$$

Proof. The proof is postponed to Section 5.2. $\qquad\square$

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 constraining the Euclidean component X to stay in the domain **D** by making it jump when it reaches the boundary **B**. The Euclidean component becomes a **D**-valued piecewise continuous and deterministic process with only forced jumps when it hits the boundary **B**. The Y-component is, as before, a **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

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

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

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

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 $n \in \mathbb{N}$. Then we can define a random variable σ_n, the inter-jump time, satisfying

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

We define the nth 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 nth 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 **D** is given by

$$\nu_y(\mathrm{d}w \mid \tilde{X}_n)\mathbf{1}_{\mathbf{B}}(\tilde{X}_n) + \delta_{\tilde{X}_n}(\mathrm{d}w)\mathbf{1}_{\mathbf{B}^c}(\tilde{X}_n)$$

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

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

That is to say, 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 it reaches 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 $i \in \mathbb{N}$,

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

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

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = F_{Y(t)}(X(t)), \quad t \in \left[T_{i-1}^{\mathbf{B}}, T_i^{\mathbf{B}}\right), \; i \in \mathbb{N}.$$

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

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

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

Proposition 5. *Suppose that Assumption* A *holds and that*

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

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

Proof. See [7, 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 **D** far enough from **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 $\nu_y(\cdot \mid x)$ has support outside the semi-tube $\mathcal{T}_\delta(\mathbf{B}) \cap \mathbf{D}$. Then, in this case, the number of times the process can reach the boundary **B** is at most $\lfloor T\delta^{-1} \sup_{(x,y) \in \mathbf{D} \times \mathbf{Y}} \|F_y(x)\| \rfloor$, 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 D and D.2 hold. Recall that the averaged field is defined, for $x \in \overline{\mathbf{D}}$, by

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

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

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

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 $\nu_y(\cdot \mid x)$ 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 $\mathbb{R}^d$-valued processes defined on $[0, T]$, that is, they belong to the Skorokhod space $\mathbf{D}([0, T], \mathbb{R}^d)$. We consider this space endowed with the so-called J_1 topology; see [16] 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 $\overline{X}_0$, 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 $(\overline{X}(t))_{t\in[0,T]}$ starting at $\overline{X}_0$ and the following hold.*

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

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

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

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{X}(t) = \overline{F}(\overline{X}(t)),$$

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

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

Proof. The proof is postponed to Section 5.3. $\qquad\square$

4. Examples and possible extensions

We start this section in Section 4.1 with an academic example showing what kind of situation our main result can be applied to. We then go on to discuss some possible extensions of our main result in Section 4.2. In Section 4.2.3 we discuss the application of our main result to the reduction of a conductance-based neuron model.

4.1. A constrained centrifugal motion in dimension two

Consider the unit circle

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

with the unit open disc

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

as corresponding domain, and outward normal $n_{\mathbf{B}}(x) = x$. For a finite set of angles $\mathbf{Y}$ in $(-\pi/2, \pi/2)$, such as

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

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

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

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

$$F_y(x) := \lambda R_y x \quad \text{for all } x \in \overline{\mathbf{D}}$$

with $\lambda > 0$. These vector fields satisfy Assumption A.1. Due to the fact that $\mathbf{Y} \subset (-\pi/2, \pi/2)$, 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

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

These 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 $\nu_y(\mathrm{d}w \mid x)$ at the boundary, has the 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). By change of variables, we have

$$\nu_y(\mathrm{d}w \mid x) = f_U(\psi_1(w \mid x)) f_{\Theta_y}(\psi_2(w \mid x)) \mid \det J_\psi(w \mid x) \mid \mathrm{d}w$$

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

$$\psi(w \mid x) := (\psi_1(w \mid x), \psi_2(w \mid x))$$

$$:= \left(1 - \frac{1}{2} \frac{\|x - w\|^2}{1 - x \cdot w}, \tan^{-1}\left(\frac{x \cdot w^\perp}{1 - x \cdot w} \right) \right)$$

with

$$w^\perp = \begin{pmatrix} -w_2 \\ w_1 \end{pmatrix}.$$

Note that if w satisfies $x \cdot w = 1$, then w is not in the support of $\nu_y(\cdot \mid x)$. 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 this is continuous, we can apply Theorem 4. For $x \in \overline{\mathbf{D}}$, the average vector field is given by

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

and the averaged measure at the boundary by

$$\overline{\nu}(\mathrm{d}w \mid x) = \sum_{y \in \mathbf{Y}} f_{\Theta_y}(\psi_2(w \mid x)) \mu_y^{\mathbf{B}}(x) f_U(\psi_1(w \mid x)) \mid \det J_\psi(w \mid x) \mid \mathrm{d}w,$$

with

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

which is independent of x. Two trajectories of the process are displayed in Figure 2, 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. To show them we require only minor modifications of the convergence proof presented in Section 5.3.
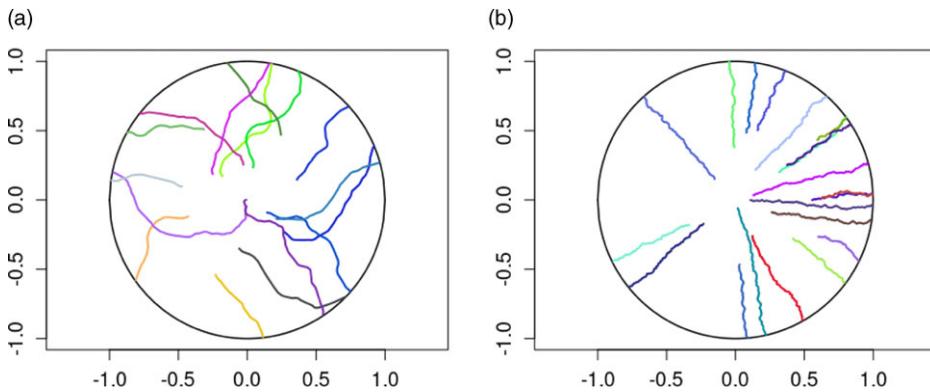
FIGURE 2. (a) Trajectory of the process with $\lambda = 0.1$, $n = 10$ and $\varepsilon = 1$ up to the horizon time $T = 300$. The initial condition is distributed as a pair of two independent Beta distributions supported by $[-1/2, 1/2]$ with parameters $(2, 2)$ for the X-component and is equal to 5 for the Y-component. We coloured each piece of the trajectory between two forced jumps with a different colour. (b) The same but with $\varepsilon = 0.001$. For the simulation of piecewise deterministic Markov processes, we refer to [26].

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 $\mathbb{R}^d$ with a 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

$$\nu_{Y(T_1^{\mathbf{B},-})}\big(\cdot \mid X(T_1^{\mathbf{B},-})\big)$$

is now valued in the domain $\mathbf{D}_1$. This should be emphasized by writing

$$\nu_{1, Y(T_1^{\mathbf{B},-})}\big(\cdot \mid X(T_1^{\mathbf{B},-})\big).$$

And so on and so forth. Our assumptions, in particular 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*

$$\overline{F}(x) \cdot n_{\mathbf{B}_i}(x) > 0 \quad \text{for all } x \in \mathbf{B}_i.$$

We could also change the form of the flow after each forced jump by indexing the vector fields via the successive domains. This change of domain framework is the one considered in [7] for the definition of a general PDMP.

4.2.2. *Slow–fast discrete process.* Another natural framework is the one where the discrete component itself exhibits 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 $Q_j^{(\mathrm{F})}$ of the transition rate matrix $Q^{(\mathrm{F})}$; see [29, Section 3.6].

This multi-scale framework is a natural fit for some applications, such as conductance-based neuron models in neuroscience [22]. This framework is fully developed in [29] in the case of continuous-time Markov chains, where other kinds of applications are considered.

In contrast to the extension presented in the previous section, Assumption B requires further reformulation in order to deal with this multi-scale case. We follow [22].

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 $\mathbf{Y}_i$ which is the solution of $\pi^{(i)}(x)Q_i^{(\mathrm{F})}(x) = 0$.*

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

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

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

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

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

$$\overline{F}_i(x) \cdot n_{\mathbf{B}}(x) > 0 \quad \text{for all } x \in \mathbf{B}.$$

Then, when $\overline{X}$ 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

$$\overline{v}_i(\cdot \mid x) = \frac{\sum_{y \in \mathbf{Y}_i} v_y(\cdot \mid x)\pi_y^{(i)}(x)(F_y(x) \cdot n_{\mathbf{B}}(x))^+}{\sum_{y \in \mathbf{Y}_i} \pi_y^{(i)}(x)(F_y(x) \cdot n_{\mathbf{B}}(x))^+}.$$

4.2.3. *Reduction of conductance-based neuron models: the case of the Morris–Lecar model.* The Morris–Lecar model is a conductance-based biological neuron model for the generation of an action potential that was proposed in 1981 [21] and has continued to be used and studied to this day; see for example [2]. We consider a hybrid version of this model, where the membrane potential of the neuron, denoted by v, and the proportion of open potassium ion channels, denoted by w, follow a differential equation given by

$$\frac{\mathrm{d}}{\mathrm{d}s}x(s) = F_y(x(s)), \quad s \in [0, T],$$

with

$$x = \begin{pmatrix} v \\ w \\ t \end{pmatrix}$$

and

$$F_y(x) = \begin{pmatrix} \frac{1}{C}(I - g_{\text{Ca}}y(v - v_{\text{Ca}}) - g_{\text{K}}w(v - v_{\text{K}}) - g_{\text{L}}(v - v_{\text{L}})) \\ (1 - w)\alpha_{\text{K}}(v) - \beta_{\text{K}}(v)w \\ 1 \end{pmatrix}.$$

The definitions of the various functions and constants that appear in the model are given in Appendix A. The reason for including time t explicitly as a variable should become clear later on. The variable y accounts for the proportion of open calcium channels throughout the neuronal membrane. We assume, in this hybrid version of the model, that

$$y = \frac{1}{N}\sum_{i=1}^{N} y_i,$$

where, for an integer $N \geq 2$ and $i \in \{1, \ldots, N\}$, $y_i \in \{0, 1\}$ is the current state of a process Y_i that jumps between 0 and 1 at v-dependent rates $\alpha_{\text{Ca}}(v)$ and $\beta_{\text{Ca}}(v)$ given in Appendix A. Thus the state space for the mode-component of the process is

$$\mathbf{Y} = \left\{ \frac{k}{N}; k \in \{0, \ldots, N\} \right\}.$$

It is very common (see [21], [22]) to take advantage of the different speeds of the dynamics of v and w on the one hand, and of y on the other hand, in order to simplify the Morris–Lecar model. Indeed, the ratio between the time-scales of w and y is of the order of a tenth. It is therefore natural to accelerate the dynamic of y to obtain a reduced model, perhaps easier to study. Another way of simplifying, also very common in theoretical neuroscience, is to simplify the dynamics of the action potential by considering that when v reaches a certain threshold v_{th}, an action potential is triggered: the variables v and w are then instantly brought back to some resting levels v_{rest} and w_{rest}. Moreover, a refractory period is often added, preventing the neuron from spiking before some possibly random time t_r in order to mimic the qualitative properties of real neurons. These mechanisms correspond to that of the integral and fire models, for which the reference [4] can be consulted.

In this setting and with appropriate initial conditions, the variable $x = (v\, w\, t)^T$ evolves in a physiological domain which is a rectangular box $[v_-, v_{\text{th}}] \times [0, 1] \times [0, T]$; see [27, Section 3.4.3]. Of course, this box does not have a $\mathcal{C}^1$-boundary. In order to apply Theorem 4 and obtain an appropriate reduced model, we need to consider an auxiliary $\mathcal{C}^1$-boundary $\mathbf{B}$ whose projection on the $(v\, w)$-phase space coincides with the rectangle $[v_-, v_{\text{th}}] \times [0, 1]$ on $\{v_{\text{th}}\} \times [0, 1]$, such as the one represented in Figure 3. The jump measure at the boundary is given, for any $x = (v\, w\, t)^T \in \mathbf{B}$ and $y \in \mathbf{Y}$, by

$$\nu_y(\mathrm{d}\tilde{x} \mid x) = f(\tilde{x} \mid x)\, \mathrm{d}\tilde{x},$$

where the density f is given in Appendix A. We have depicted the projection of the support of this measure in the $(v\, w)$-phase space in Figure 3. The interested reader can check that Assumptions A, B, C, D.2, and E are fulfilled for the described model. Applying Theorem 4, the reduced and constrained Morris–Lecar model follows the dynamic given by
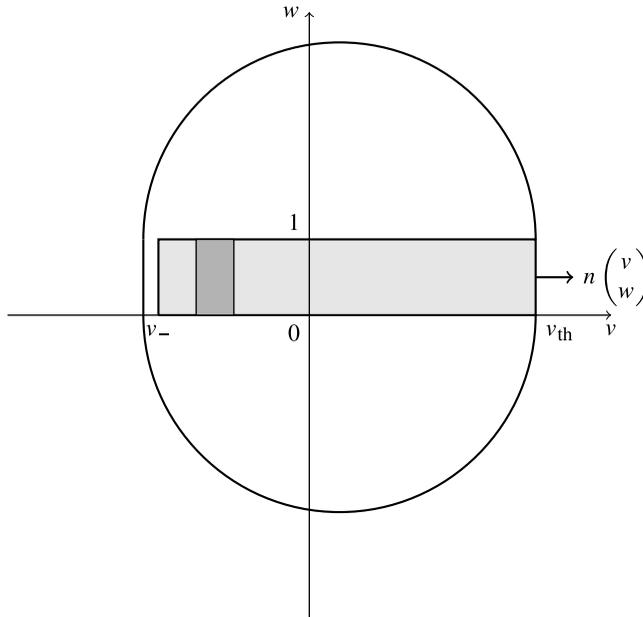
FIGURE 3. The projection of the physiological domain on the $(v\,w)$-phase space is shown in light grey. The support of the corresponding jump measure is in dark grey. The projection of a possible auxiliary mathematical domain **B** is indicated by bold black lines. An instance of the normal vector field $n\,(v\,w)$ is also represented.

the averaged vector field

$$
\overline{F}(x) = \begin{pmatrix} \frac{1}{C}(I - g_{Ca}y_{\infty}(v)(v - v_{Ca}) - g_{K}w(v - v_{K}) - g_{L}(v - v_{L})) \\ (1 - w)\alpha_{K}(v) - \beta_{K}(v)w \\ 1 \end{pmatrix}
$$

where

$$
y_{\infty}(v) = \frac{\alpha_{Ca}(v)}{\alpha_{Ca}(v) + \beta_{Ca}(v)}.
$$

As for the averaged jump measure, it remains unchanged: $\overline{\nu}(\mathrm{d}\tilde{x} \mid x) = f(\tilde{x} \mid x)\,\mathrm{d}\tilde{x}$. In Figure 4 we display the course of the action potential through the simulations of the Morris–Lecar model in its two time-scale and averaged versions. We also display this course for the deterministic Morris–Lecar model for comparison. The numerical constants used for the simulations are given in Appendix A. Of course, the hybrid and constrained models, for both the multiscale and averaged versions, differ qualitatively from the deterministic model. However, an important part of the information carried out by a neural cell lies in the distribution of the interspike intervals or in the distribution of the time to first spike; see [13]. In this context, the hybrid and constrained as well as the reduced models are certainly interesting to study.
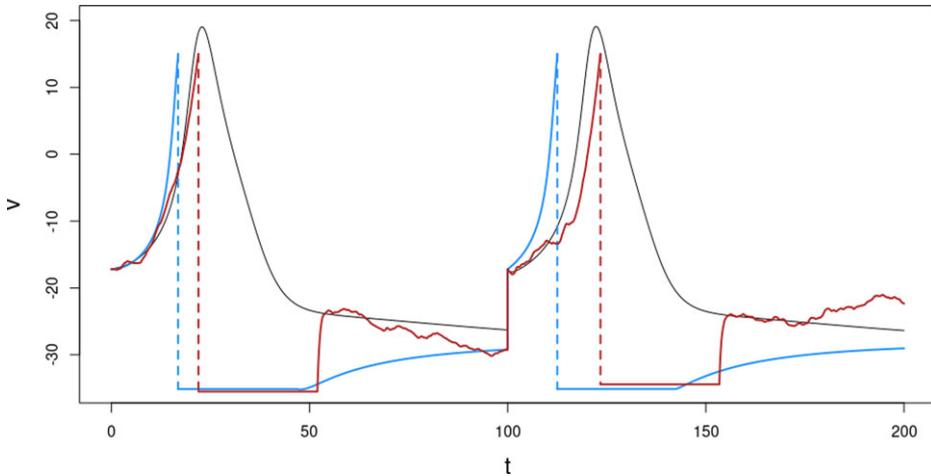
FIGURE 4. Course of the action potential (in mV) against time (in ms) according to three different versions of the Morris–Lecar model. Thin black lines indicate the course of the action potential for the deterministic Morris–Lecar model; see [21]. Red lines show this course for the hybrid and constrained model with two time-scales ($\varepsilon = 1$). Blue lines show the averaged version of the latter model. Two pulses, at time 0 and 100 (in ms), have been injected into the system in order to induce two action potentials by shifting the current up to -17.2 mV at these time points.

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

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

since under Assumption D.3, the process X_ε may cross the hypersurface at most once. 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 (5). To get the formula (6) in the proposition, we need to show that for any measurable and bounded function h and for any $\varepsilon > 0$, we have

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

For this purpose, note that under the stated assumptions (see [3])

$$\mathbf{E}\big(h\big(X_\varepsilon\big(T_\varepsilon^{\mathbf{B}}\big), Y_\varepsilon\big(T_\varepsilon^{\mathbf{B}}\big)\big)\mathbf{1}_{[0,T]}\big(T_\varepsilon^{\mathbf{B}}\big)\big)$$

$$= \mathbf{E}\bigg(\lim_{\delta \to 0} \frac{1}{2\delta} \int_0^T h(X_\varepsilon(s), Y_\varepsilon(s))F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot n_B(X_\varepsilon(s))\|\nabla\rho(X_\varepsilon(s))\|$$

$$\mathbf{1}_{[-\delta,\delta]}(\rho(X_\varepsilon(s)))\, \mathrm{d}s\bigg).$$

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

$$\frac{1}{2\delta} \int_0^T h(X_\varepsilon(s), Y_\varepsilon(s)) F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot n_B(X_\varepsilon(s)) \|\nabla\rho(X_\varepsilon(s))\| \mathbf{1}_{[-\delta,\delta]}(\rho(X_\varepsilon(s))) \, \mathrm{d}s$$

$$\leq \sup_{(x,y)\in\mathbb{R}^d\times\mathbf{Y}} |h(x,y)|,$$

the dominated convergence theorem yields

$$\mathbf{E}\big(h\big(X_\varepsilon\big(T_\varepsilon^{\mathbf{B}}\big), Y_\varepsilon\big(T_\varepsilon^{\mathbf{B}}\big)\big)\mathbf{1}_{[0,T]}\big(T_\varepsilon^{\mathbf{B}}\big)\big)$$

$$= \lim_{\delta\to 0} \mathbf{E}\bigg(\frac{1}{2\delta} \int_0^T h(X_\varepsilon(s), Y_\varepsilon(s)) F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot n_B(X_\varepsilon(s)) \|\nabla\rho(X_\varepsilon(s))\|$$

$$\mathbf{1}_{[-\delta,\delta]}(\rho(X_\varepsilon(s))) \, \mathrm{d}s\bigg).$$

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

$$\mathbf{E}\big(h\big(X_\varepsilon\big(T_\varepsilon^{\mathbf{B}}\big), Y_\varepsilon\big(T_\varepsilon^{\mathbf{B}}\big)\big)\mathbf{1}_{[0,T]}\big(T_\varepsilon^{\mathbf{B}}\big)\big)$$

$$= \lim_{\delta\to 0} \frac{1}{2\delta} \int_{\mathbb{R}^d} \sum_{y\in\mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \int_0^T f_{s,\varepsilon}(x, y) \, \mathrm{d}s \|\nabla\rho(x)\| \mathbf{1}_{[-\delta,\delta]}(\rho(x)) \, \mathrm{d}x.$$

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

$$\lim_{\delta\to 0} \frac{1}{2\delta} \int_{\mathbb{R}^d} \sum_{y\in\mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \int_0^T f_{s,\varepsilon}(x, y) \, \mathrm{d}s \|\nabla\rho(x)\| \mathbf{1}_{[-\delta,\delta]}(\rho(x)) \, \mathrm{d}x$$

$$= \int_{\mathbf{B}} \sum_{y\in\mathbf{Y}} h(x, y) F_y(x) \cdot n_{\mathbf{B}}(x) \int_0^T f_{s,\varepsilon}(x, y) \, \mathrm{d}s \sigma_{\mathbf{B}}(\mathrm{d}x).$$

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

5.2. Proof of Proposition 4

For the proof of Proposition 3 to still be valid when Assumption D.3 is replaced 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 Theorem 2 and Skorokhod's representation theorem, one can assume that (X_{ε_k}) converges to $\overline{X}$ 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 $\overline{X}$ crosses $\mathbf{B}$, and does it exactly once by Assumption D.2. Note that this also implies, for all sufficiently large k, 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 $T_{\mathrm{down},\varepsilon_k}$ of $\mathbf{B}$ and the last up-crossing $T_{\mathrm{last},\varepsilon_k}$, the increasing and decreasing phases of $\rho(X_{\varepsilon_k}(\cdot))$ must compensate.

That is to say, for k large enough,

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

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

We then have, for k large enough,

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

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

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

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

$$\mathbf{E}\left(\int_{\mathbf{B}}\overline{F}(x)\cdot n_{\mathbf{B}}(x)\overline{f}_{\overline{T}^{\mathbf{B}}}(x)\sigma_{\mathbf{B}}(\mathrm{d}x)\right)=0. \tag{7}$$

Under Assumption D.2 we have

$$\overline{F}(x)\cdot n_{\mathbf{B}}(x)>0\quad\text{for all }x\in\mathbf{B}.$$

Moreover, since $\overline{X}$ crosses $\mathbf{B}$, we also have $f_{\overline{T}^{\mathbf{B}}}(x)>0$ in some neighbourhood of $\mathbf{B}$. These two facts are in contradiction to equation (7).

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 programme initiated by Prokhorov (see [16]): in Section 5.3.1 we prove that the family $\{X_\varepsilon,\ \varepsilon\in(0,1]\}$ is tight in $\mathbf{D}([0,T],\mathbb{R}^d)$, and in Section 5.3.2 we identify the limit.

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 described in [11] and [20]. 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 us 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 [18, 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

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

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 A and E, there is at most one forced jump between times t and $t + h$ for h small enough. Indeed, Assumptions A and E imply that at any time, the time needed to reach the boundary is bounded from below by a strictly positive time, independently of X_ε, Y_ε, and ε. This implies that we can take h small enough so that we can reach the boundary at most once in a time window of size h. 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, [14, 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 $t \in [0, T]$ that is uniformly bounded in ε. Therefore

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

uniformly in $\varepsilon \in (0, 1]$, where the definition of the tube $\mathcal{T}_{\sup_{(x,y) \in \mathbf{D} \times \mathbf{Y}} \|F_y(x)\| h}(\mathbf{B})$ is given in Section 3.1. The family $\{X_\varepsilon; \, \varepsilon \in (0, 1]\}$ is thus tight in $\mathbf{D}([0, T], \mathbb{R}^d)$, endowed with the Skorokhod topology.

5.3.2. *Identification of the limit.* Let $\overline{X}$ be an accumulation point of the tight family $\{X_\varepsilon; \, \varepsilon \in (0, 1]\}$. We consider a real-valued and measurable function h on $\overline{\mathbf{D}}$ such that the following hold.

- The functions $t \mapsto h(\phi_y(x, t))$ are absolutely continuous on $[0, t^{\mathbf{B}}(x, y)[$ for any $(x, y) \in \overline{\mathbf{D}} \times \mathbf{Y}$.

- For any $\varepsilon > 0$, there is some sequence of stopping time (t_n) going to infinity such that

$$\mathbf{E}\left(\sum_{s \leq T} |h(X_\varepsilon(s \wedge t_n)) - h(X_\varepsilon(s \wedge t_n^-))| \right) < \infty,$$

the same being true if we replace X_ε with $\overline{X}$.

Our aim is to show that we have

$$\mathbf{E}(h(\overline{X}(t))) = \mathbf{E}(h(\overline{X}(0))) + \mathbf{E}\left(\int_0^t \overline{F}(\overline{X}(s)) \cdot \nabla h(\overline{X}(s)) \, \mathrm{d}s \right)$$

$$+ \mathbf{E}\left(\int_0^t \int_{\mathbf{D}} \left[h(w) - h(\overline{X}(s^-)) \right] \overline{v}(\mathrm{d}w \mid \overline{X}(s^-)) \overline{p}^{\mathbf{B}}(\mathrm{d}s) \right),$$

for any $t \in [0, T]$, as this will characterize the law of $\overline{X}$, as stated in [7, Theorem 26.14]. We will still let $(X_\varepsilon)_\varepsilon$ denote a sequence converging towards $\overline{X}$ in the Skorokhod topology. By convergence in the Skorokhod topology, we have, for any point of continuity $t \in [0, T]$ of $\overline{X}$,

$$\mathbf{E}(h(\overline{X}(t))) = \lim_{\varepsilon \to 0} \mathbf{E}(h(X_\varepsilon(t))).$$

Following [7, Theorem 26.14], we also know that for any $\varepsilon \in (0, 1]$ and $t \in [0, T]$,

$$\mathbf{E}(h(X_\varepsilon(t))) = \mathbf{E}(h(X_\varepsilon(0))) + \mathbf{E}\left(\int_0^t F_{Y_\varepsilon(s)}(X_\varepsilon(s)) \cdot \nabla h(X_\varepsilon(s)) \, \mathrm{d}s \right)$$

$$+ \mathbf{E}\left(\int_0^t \int_{\mathbf{D}} [h(w) - h(X_\varepsilon(s^-))] v_{Y_\varepsilon(s^-)}(\mathrm{d}w \mid x_\varepsilon(s^-)) \overline{p}_\varepsilon^{\mathbf{B}}(\mathrm{d}s) \right).$$

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

Proposition 6. *We have*

$$\lim_{\varepsilon \to 0} \mathbf{E}(h(X_\varepsilon(0))) = \mathbf{E}(h(\overline{X}(0))),$$

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

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

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 expression considered, this is a direct consequence of the convergence in the Skorokhod topology of (X_ε) towards $\overline{X}$ and of standard averaging techniques for piecewise deterministic Markov processes without forced jumps, as presented in [22]. □

It remains to show that

$$\lim_{\varepsilon \to 0} \mathbf{E}\left(\int_0^t \int_{\mathbf{D}} [h(w) - h(X_\varepsilon(s^-))] \nu_{Y_\varepsilon(s^-)}(dw \mid x_\varepsilon(s^-)) \overline{p}_\varepsilon^{\mathbf{B}}(ds) \right)$$

$$= \mathbf{E}\left(\int_0^t \int_{\mathbf{D}} [h(w) - h(\overline{X}(s^-))] \overline{\nu}(dw \mid \overline{X}(s^-)) \overline{p}^{\mathbf{B}}(ds) \right)$$

Since we cannot 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 $t \in [0, T]$,

$$\mathbf{E}\left(\int_0^t \int_{\mathbf{D}} [h(w) - h(X_\varepsilon(s^-))] \nu_{Y_\varepsilon(s^-)}(dw \mid X_\varepsilon(s^-)) \overline{p}_\varepsilon^{\mathbf{B}}(ds) \right)$$

$$= \sum_{i=1}^{\infty} \mathbf{E}\big(\Gamma\big(X_\varepsilon\big(T_{i,\varepsilon}^{\mathbf{B},-}\big), Y_\varepsilon\big(T_{i,\varepsilon}^{\mathbf{B},-}\big)\big) \mathbf{1}_{T_{i,\varepsilon}^{\mathbf{B}} \leq t} \big),$$

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

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

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

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

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

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

where $\mu^{\mathbf{B}}$ is defined in Definition 3. The proof will rest 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 jumping after the ith collision of the boundary.

2. These companion processes allow us to make the link with the second main ingredient: 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 $i \in \mathbb{N}$ and $\varepsilon \in (0, 1]$, we let $(V_\varepsilon^{(i)}, W_\varepsilon^{(i)})$ denote the process constructed as $(X_\varepsilon, Y_\varepsilon)$ until the ith collision of the boundary by $V_\varepsilon^{(i)}$ at time $T_{i,\varepsilon}^{(i),\mathbf{B}}$. Then, if $T_{i,\varepsilon}^{(i),\mathbf{B}}$ is finite, for $t \geq T_{i,\varepsilon}^{(i),\mathbf{B}}$,

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

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

In particular, for $V_\varepsilon^{(i)}$, there is no forced jump after the $(i-1)$th forced jump, the process $V_\varepsilon^{(i)}$ remaining continuous after this time. Usual averaging results allow us to consider, for example, the limit of the first companion process.

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

$$\overline{V}^{(1)}(t) = \overline{X}_0 + \int_0^t \overline{F}(\overline{V}^{(1)}(s)) \, \mathrm{d}s,$$

where $\overline{F}$ *is given in Definition 3.*

Proof. Under the assumptions made, the result is given by Theorem 2. $\square$

Remark 9. When $i \geq 2$, the same result holds true for the processes $(V_\varepsilon^{(i)})_{\varepsilon \in (0,1]}$ when far from the boundary. Indeed, for $i \geq 2$, we can show that the family $(V_\varepsilon^{(i)})_{\varepsilon \in (0,1]}$ is tight in $\mathbf{D}([0, T], \mathbb{R}^d)$ 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,

$$\frac{\mathrm{d}}{\mathrm{d}t} \overline{V}^{(i)}(t) = \overline{F}(\overline{V}^{(i)}(t)).$$

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

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

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

Proposition 8. *We have*

$$\lim_{\varepsilon \to 0} \mathbf{E}\big(\Gamma\big(X_\varepsilon\big(T_{1,\varepsilon}^{\mathbf{B},-}\big), Y_\varepsilon\big(T_{1,\varepsilon}^{\mathbf{B},-}\big)\big)\mathbf{1}_{T_{1,\varepsilon}^{\mathbf{B}} \leq t}\big) = \mathbf{E}\big(\overline{\Gamma}\big(\overline{X}\big(\overline{T}_1^{\mathbf{B},-}\big)\big)\mathbf{1}_{\overline{T}_1^{\mathbf{B}} \leq t}\big)$$

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

$$\lim_{\varepsilon \to 0} \mathbf{P}(T_{1,\varepsilon}^{\mathbf{B},-} \le t) = \mathbf{P}(\overline{T}_1^{\mathbf{B},-} \le t).$$

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

Proposition 9. *For* $i = 2$,

$$\lim_{\varepsilon \to 0} \mathbf{E}\big(\Gamma\big(X_\varepsilon\big(T_{2,\varepsilon}^{\mathbf{B},-}\big), Y_\varepsilon\big(T_{2,\varepsilon}^{\mathbf{B},-}\big)\big)\mathbf{1}_{T_{2,\varepsilon}^{\mathbf{B}} \le t}\big) = \mathbf{E}\big(\overline{\Gamma}\big(\overline{X}\big(\overline{T}_2^{\mathbf{B},-}\big)\big)\mathbf{1}_{\overline{T}_2^{\mathbf{B}} \le t}\big).$$

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

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

where $T_{1,\varepsilon}^{\mathbf{B},(2)}$ is the time of the first forced jump for the second companion process. It has the 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

$$\mathbf{E}\big(N_{\mathbf{B},y}\big(\big[T_{1,\varepsilon}^{\mathbf{B},-,(2)}, t\big], V_\varepsilon^{(2)}\big)\big) = \mathbf{E}\big(\mathbf{E}\big(N_{\mathbf{B},y}\big(\big[T_{1,\varepsilon}^{\mathbf{B},(2),-}, t\big], V_\varepsilon^{(2)}\big) \mid \mathcal{F}_{T_{1,\varepsilon}^{\mathbf{B},(2),-}}\big)\mathbf{1}_{T_{1,\varepsilon}^{\mathbf{B},(2)} \le t}\big)$$

$$= \int_0^t \mathbf{E}\big(N_{\mathbf{B},y}\big([s, t], V_\varepsilon^{(2)}\big)\big)\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{P}\big(T_{1,\varepsilon}^{\mathbf{B},(2)} \le s\big)\,\mathrm{d}s$$

$$= \int_0^t \mathbf{E}\big(N_{\mathbf{B},y}\big([s, t], V_\varepsilon^{(2)}\big)\big)\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{P}\big(T_{1,\varepsilon}^{\mathbf{B},(1)} \le s\big)\,\mathrm{d}s,$$

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

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

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

$$\lim_{\varepsilon \to 0} \mathbf{E}\big(N_{\mathbf{B},y}\big(\big[T_{1,\varepsilon}^{\mathbf{B},-}, t\big], V_\varepsilon^{(2)}\big)\big)$$

$$= \int_0^t \int_{\mathbf{B}} (F_y(x) \cdot n_{\mathbf{B}}(x))^+ \pi_y(x) \int_s^t f_{\overline{X}(u)}(x)\,\mathrm{d}u\sigma^{\mathbf{B}}(\mathrm{d}x)\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{P}\big(\overline{T}_1^{\mathbf{B}} \le s\big)\,\mathrm{d}s.$$

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

$$\lim_{\varepsilon \to 0} \mathbf{E}\big(\Gamma\big(X_\varepsilon\big(T_{2,\varepsilon}^{\mathbf{B},-}\big), Y_\varepsilon\big(T_{2,\varepsilon}^{\mathbf{B},-}\big)\big)\mathbf{1}_{T_{2,\varepsilon}^{\mathbf{B}} \le t}\big)$$

$$= \int_0^t \int_{\mathbf{B}} \sum_{y \in \mathbf{Y}} \Gamma(x, y)(F_y(x) \cdot n_{\mathbf{B}}(x))^+ \pi_y(x) \int_s^t f_{\overline{X}(u)}(x)\,\mathrm{d}u\sigma^{\mathbf{B}}(\mathrm{d}x)\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{P}\big(\overline{T}_1^{\mathbf{B}} \le s\big)\,\mathrm{d}s.$$

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

TABLE 1. Parameters for numerical integration and stochastic simulation.

v_1	v_2	v_3	v_4	g_{Ca}	g_K	g_L
0 mV	15 mV	10 mV	10 mV	4 mS cm^{-2}	8 mS cm^{-2}	2 mS cm^{-2}
v_{Ca}	v_K	v_L	C	λ	w_{rest}	v_{rest}
100 mV	−70 mV	−50 mV	20 mF cm^{-2}	0.1	0.022314	−35 mV

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

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

The proof of Theorem 4 is thus complete.

Appendix A. Functions and constants for the Morris–Lecar model

The auxiliary functions of the Morris–Lecar model [21] used in the present article are

$$\alpha_{Ca}(v) = \frac{1}{2}\lambda \cosh\left(\frac{v - v_1}{2v_2}\right)\left(1 + \tanh\left(\frac{v - v_1}{v_2}\right)\right),$$

$$\alpha_K(v) = \frac{1}{2}\cosh\left(\frac{v - v_3}{2v_4}\right)\left(1 + \tanh\left(\frac{v - v_3}{v_4}\right)\right),$$

$$\beta_{Ca}(v) = \frac{1}{2}\lambda \cosh\left(\frac{v - v_1}{2v_2}\right)\left(1 - \tanh\left(\frac{v - v_1}{v_2}\right)\right),$$

$$\beta_K(v) = \frac{1}{2}\cosh\left(\frac{v - v_3}{2v_4}\right)\left(1 - \tanh\left(\frac{v - v_3}{v_4}\right)\right).$$

The values of the parameters and initial conditions used for the numerical integration and stochastic simulations are given in Table 1. The initial potential is $v_0 = -17.2$ mV and the initial proportion of open potassium channels is $w_0 = 0.022314$. For the hybrid model, the number of calcium ion channels has been taken to be $N = 100$. The initial number of open channels is 10. For the constrained models, we have set $v_{th} = 15$.

The density of the jump measure at the boundary is given, for any $x = (v\ w\ t)^T \in \mathbf{B}$, $\tilde{x} = (\tilde{v}\ \tilde{w}\ \tilde{t})^T \in \mathbf{D}$, by

$$f(\tilde{x} \mid x) = f_1(\tilde{v})f_2(\tilde{w})f_3(\tilde{t} \mid t),$$

where f_1 and f_2 are the densities of Beta laws supported by $[-38, -32]$ and $[w_{rest} - 0.1, w_{rest} + 0.1]$ respectively, while $f_3(\cdot \mid t)$ is the density of a Beta distribution supported by $[t + t_r - 10, t + t_r + 10]$ with $t_r = 30$ ms.

Acknowledgements

Funding information

There are no funding bodies to thank relating to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] AZAÏS, R. AND GÉNADOT, A. (2019). Estimation of the average number of continuous crossings for non-stationary non-diffusion processes. *J. Statist. Planning Infer.* **198**, 119–138.

[2] BAO, H., YU, X., XU, Q., WU, H. AND BAO, B. (2022). Three-dimensional memristive Morris–Lecar model with magnetic induction effects and its FPGA implementation. *Cogn. Neurodynam.* Available at https://doi.org/10.1007/s11571-022-09871-6.

[3] BOROVKOV, K. AND LAST, G. (2012). On Rice's formula for stationary multivariate piecewise smooth processes. *J. Appl. Prob.* **49**, 351–363.

[4] BURKITT, A. N. (2006). A review of the integrate-and-fire neuron model, I: Homogeneous synaptic input. *Biol. Cybernet.* **95**, 1–19.

[5] COSTANTINI, C. AND KURTZ, T. (2015). Viscosity methods giving uniqueness for martingale problems. *Electron. J. Prob.* **20**, 1–27.

[6] COSTANTINI, C. AND KURTZ, T. G. (2018). Existence and uniqueness of reflecting diffusions in cusps. *Electron. J. Prob.* **23**, 1–21.

[7] DAVIS, M. H. A. (1993). *Markov Models & Optimization* (Monographs on Statistics and Applied Probability **49**). CRC Press.

[8] ENGEL, K.-J. AND NAGEL, R. (2001). One-parameter semigroups for linear evolution equations. *Semigroup Forum* **63**, 278–280.

[9] FAGGIONATO, A., GABRIELLI, D. AND RIBEZZI CRIVELLARI M. (2010). Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors. *Markov Process. Relat. Fields* **16**, 497–548.

[10] GÉNADOT, A. (2013). A multiscale study of stochastic spatially-extended conductance-based models for excitable systems. Doctoral thesis, Université Pierre et Marie Curie-Paris VI.

[11] GÉNADOT. (2019). Averaging for some simple constrained Markov processes. *Prob. Math. Statist.* **39**, 139–158.

[12] GERSTNER, W., KISTLER, W. M., NAUD, R. AND PANINSKI, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press.

[13] GOLLISCH, T. AND MEISTER, M. (2008). Rapid neural coding in the retina with relative spike latencies. *Science* **319**, 1108–1111.

[14] GWIŻDŻ, P. AND TYRAN-KAMIŃSKA, M. (2019). Densities for piecewise deterministic Markov processes with boundary. *J. Math. Anal. Appl.* **479**, 384–425.

[15] HAIRER, M. AND LI, X.-M. (2020) Averaging dynamics driven by fractional Brownian motion. *Ann. Prob.* **48**, 1826–1860.

[16] JAKUBOWSKI, A. (2007). The Skorokhod space in functional convergence: a short introduction. In *International Conference: Skorokhod Space*, pp. 11–18.

[17] KAC, M. (1943). On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.* **49**, 314–320.

[18] KALLENBERG, O. (1997). *Foundations of Modern Probability*. Springer, New York.

[19] KRANTZ, S. AND PARKS, H. (1999). *The Geometry of Domains in Space*. Springer.

[20] KURTZ, T. G. (1990). Martingale problems for constrained Markov processes. In *Recent Advances in Stochastic Calculus*, ed. J. S. Baras and V. Mirelli, pp. 151–168. Springer.

[21] MORRIS, C. AND LECAR, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* **35**, 193–213.

[22] PAKDAMAN, K., THIEULLEN, M. AND WAINRIB, G. (2012). Asymptotic expansion and central limit theorem for multiscale piecewise-deterministic Markov processes. *Stoch. Process. Appl.* **122**, 2292–2318.

[23] PARDOUX, E. AND VERETENNIKOV, A. Y. (2001). On the Poisson equation and diffusion approximation, I. *Ann. Prob.* **29**, 1061–1085.

[24] PAVLIOTIS, G. AND STUART, A. (2008). *Multiscale Methods: Averaging and Homogenization*. Springer.

[25] PERTHAME, B. (2006). *Transport Equations in Biology*. Springer.

[26] RIEDLER, M. G. (2013). Almost sure convergence of numerical approximations for piecewise deterministic Markov processes. *J. Comput. Appl. Math.* **239**, 50–71.

[27] RIEDLER, M. G. (2011). Spatio-temporal stochastic hybrid models of biological excitable membranes. Doctoral dissertation, Heriot-Watt University.

[28] TYRAN-KAMIŃSKA, M. (2009). Substochastic semigroups and densities of piecewise deterministic Markov processes. *J. Math. Anal. Appl.* **357**, 385–402.

[29] YIN, G. G. AND ZHANG, Q. (2012). *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach*. Springer.