Piecewise deterministic Markov process - recent results

Romain Azaïs, Jean-Baptiste Bardet, Alexandre Genadot, Nathalie Krell, Pierre-André Zitt

Piecewise deterministic Markov process — recent results

Romain Azaïs* Jean-Baptiste Bardet† Alexandre Génadot‡
Nathalie Krell§ Pierre-André Zitt¶

September 24, 2013

Abstract

We give a short overview of recent results on a specific class of Markov process: the Piecewise Deterministic Markov Processes (PDMPs). We first recall the definition of these processes and give some general results. On more specific cases such as the TCP model or a model of switched vector fields, better results can be proved, especially as regards long time behaviour. We continue our review with an infinite dimensional example of neuronal activity. From the statistical point of view, these models provide specific challenges: we illustrate this point with the example of the estimation of the distribution of the inter-jumping times. We conclude with a short overview on numerical methods used for simulating PDMPs.

1 General introduction

The piecewise deterministic Markov processes (denoted PDMPs) were first introduced in the literature by Davis ([Dav84, Dav93]). Already at this time, the theory of diffusions had such powerful tools as the theory of Itō calculus and stochastic differential equations at its disposal. Davis's goal was to endow the PDMP with rather general tools. The main reason for that was to provide a general framework, since up to then only very particular cases had been dealt with, which turned out not to be easily generalizable.

*Inria Bordeaux Sud-Ouest, team CQFD et Université de Bordeaux, IMB, CNRS UMR 5251, 200, Avenue de la Vieille Tour, 33405 Talence cedex, France.

†Laboratoire de Mathématiques Raphaël Salem, Université de Rouen, Avenue de l'Université, BP 12, 76801 Saint-Etienne-du-Rouvray, France.

‡Laboratoire de Probabilités et modèles aléatoires, Université Pierre et Marie Curie, 4, place Jussieu, 75005 Paris, France.

§Université de Rennes 1, Institut de Recherche mathématique de Rennes, CNRS-UMR 6625, Campus de Beaulieu. Bâtiment 22, 35042 Rennes Cedex, France.

¶Laboratoire d'Analyse et de Mathématiques Appliquées, Université de Marne-la-Vallée, 5, boulevard Descartes, Cité Descartes - Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2

PDMPs form a family of càdlàg Markov processes involving a deterministic motion punctuated by random jumps. The motion of the PDMP $\{X(t)\}_{t\geq 0}$ depends on three local characteristics, namely the jump rate λ, the flow ϕ and the transition measure Q according to which the location of the process at the jump time is chosen. The process starts from x and follows the flow $\phi(x,t)$ until the first jump time T_1 which occurs either spontaneously in a Poisson-like fashion with rate $\lambda(\phi(x,t))$ or when the flow $\phi(x,t)$ hits the boundary of the state-space. In both cases, the location of the process at the jump time T_1, denoted by $Z_1 = X(T_1)$, is selected by the transition measure $Q(\phi(x,T_1),\cdot)$ and the motion restarts from this new point as before. This fully describes a piecewise continuous trajectory for $\{X(t)\}$ with jump times $\{T_k\}$ and post jump locations $\{Z_k\}$, and which evolves according to the flow ϕ between two jumps.

These processes have been heavily studied both from a theoretical and from an applied perspective in various domains such as communication networks with the control of congestion TCP/IP (V. Dumas and al. [DGR02], V. Guillemin and al. [GRZ04]), neurobiology for the Hodgkin-Huxley model of neuronal activity (K. Pakdaman and al. [PTW10]), reliability (F. Dufour and Y. Dutuit [DD02]), biologic population models (H.G. Othmer and al.[ODA88] as well as R. Erban and H.G. Othmer [EO05]), to present just a few examples.

The paper is organized as follows. In Section 2 we give a precise definition and some general properties of the PDMPs. Then we illustrate the state of the art regarding PDMPs through three specific examples: a model of switched vector fields (Section 3), the TCP process (Section 4), and a modelization of neuronal activity (Section 5). Finally, we briefly review some results about a non-parametric statistical method to get an estimate of the conditional density associated with the jumps of a PDMP defined on a separable metric space (Section 6) and we end with a survey of numerical methods in Section 7.

2 Definition and some properties of PDMPs

2.1 Definition of a PDMP.

Let M be an open subset of $\mathbb{R}^n$, ∂M its boundary, $\overline{M}$ its closure and $\mathcal{B}(M)$ the set of real-valued, bounded, measurable functions defined on M. A PDMP is determined by its local characteristics (ϕ, λ, Q) where:

- The flow $\phi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is a one-parameter group of homeomorphisms: ϕ is continuous, $\phi(\cdot,t)$ is an homeomorphism for each $t \in \mathbb{R}$, satisfying the semigroup property: $\phi(\cdot, t+s) = \phi(\phi(\cdot,s),t)$.

  For each x in M, we introduce the deterministic hitting time of the boundary

  $$t^\star(x) := \inf\{t > 0 : \phi(x,t) \in \partial M\}, \tag{1}$$

  with the convention $\inf \emptyset = \infty$.

- The jump rate $\lambda : \overline{M} \to \mathbb{R}_+$ is assumed to be a measurable function satisfying

$$\forall x \in M,\ \exists\, \epsilon > 0 \text{ such that } \int_0^\epsilon \lambda(\phi(x,s))ds < \infty.$$

- Q is a Markov kernel on $(\overline{M}, \mathcal{B}(\overline{M}))$ satisfying the following property:

$$\forall x \in \overline{M}, \quad Q(x, M - \{x\}) = 1.$$

From these characteristics, it can be shown [Dav93, pp 62–66] that there exists a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \{\mathbb{P}_x\})$ such that the motion of the process $\{X(t)\}$ starting from a point $x \in M$ may be constructed as follows. Consider a random variable T_1 such that

$$\mathbb{P}_x\{T_1 > t\} = \begin{cases} e^{-\Lambda(x,t)} & \text{for } t < t^\star(x), \\ 0 & \text{for } t \geq t^\star(x), \end{cases}$$

where for $x \in M$ and $t \in [0, t^\star(x)]$

$$\Lambda(x,t) = \int_0^t \lambda(\phi(x,s))ds.$$

If T_1 is equal to infinity, then the process X follows the flow, i.e. $t \in \mathbb{R}_+$, $X(t) = \phi(x,t)$. Otherwise select independently a M-valued random variable (labelled Z_1) having distribution $Q(\phi(x,T_1),\cdot)$, namely $\mathbb{P}_x(Z_1 \in A) = Q(\phi(x,T_1), A)$ for any $A \in \mathcal{B}(M)$. The trajectory of $\{X(t)\}$ starting at x, for $t \in [0, T_1]$, is given by

$$X(t) = \begin{cases} \phi(x,t) & \text{for } t < T_1, \\ Z_1 & \text{for } t = T_1. \end{cases}$$

Starting from $X(T_1) = Z_1$, we now select the next inter-jump time $T_2 - T_1$ and post-jump location $X(T_2) = Z_2$ in a similar way.

This construction properly defines a Markov process $\{X(t)\}$ which satisfies the strong Markov property with jump times $\{T_k\}_{k\in\mathbb{N}}$ (where $T_0 = 0$). A very natural Markov chain is linked to $\{X(t)\}$, namely the chain $(\Theta_n)_{n\in\mathbb{N}}$ defined by $\Theta_n = (Z_n, S_n)$ with $Z_n = X(T_n)$ and $S_n = T_n - T_{n-1}$ for $n \geq 1$ and $S_0 = 0$. Clearly, the process $(\Theta_n)_{n\in\mathbb{N}}$ is a Markov chain. This chain will turn out to be particularly useful in the next sections (see Section 7).

Remark 1. Davis originally defined PDMPs on a disjoint union $\bigcup_{v\in K}\{v\} \times M_v$, where K is a countable index set and, for each v, M_v is a subset of $\mathbb{R}^n$. The definition above, with a single copy of $\mathbb{R}^n$, can be used without loss of generality – see Remark 24.9 in [Dav93] for details. Depending on the process, one or the other definition may be more natural: we will use the original definition in Section 3.

2.2 A useful discrete process

A basic assumption in all the following will be that, for every starting point $x \in M$, $\mathbb{E}\left\{\sum_k \mathbf{1}_{\{t \geq T_k\}}\right\} < \infty$. This ensures the non explosion of the process.

The chain Θ_n obtained by observing the process at jump times may not have enough jumps to guarantee good comparison properties with the continuous time process Z_t. To address this problem, O. Costa and F Dufour introduced in [DC99] another discrete chain as follows. The idea is to record the positions of the continuous process both at jump times and at additional random times given by an independent Poisson process of rate 1. Formally, first define two substochastic kernels H and J by

$$H(x, A) := \int_0^{t_\star(x)} e^{-(s+\Lambda(x,s))} \mathbf{1}_A(\phi(x, s))\, ds,$$
$$J(x, A) := \int_0^{t_\star(x)} \lambda(\phi(x, s)) e^{-(s+\Lambda(x,s))} Q(\phi(x, s), A) ds + e^{-(t_\star(x)+\Lambda(x,t_\star(x)))} Q(\phi(x, t_\star(x)), A).$$

The first kernel corresponds to the additional observation times, the second one to the "real jumps", either due to the jump rate (first term) or the hitting of the boundary (second term). The sum $G = J + H$ is proved in [DC99] to be a Markov kernel. We denote by $\{\theta_n\}$ the associated Markov chain, which can be generated from the sample paths of $\{X_t\}$ by adding observation times as said before – see [CD08, Theorem 3.1] for a formal statement.

Theorem 1 (Discrete and continuous processes, [CD08])**.** *The Markov chain $\{\theta_n\}$ and the original process $\{X_t\}$ are closely related:*

1. *The PDMP $\{X(t)\}$ is irreducible if and only if the Markov chain $\{\theta_n\}$ is irreducible.*
2. *If ν is an invariant measure for $\{X(t)\}$, then $\nu \sum_0^\infty J^j$ is invariant for $\{\theta_n\}$ and $\nu \sum_0^\infty J^j = \nu$. Conversely, if π is an invariant measure for $\{\theta_n\}$, then πH is invariant for $\{X(t)\}$ and $\pi \sum_0^\infty J^j = \pi$.*
3. *The PDMP $\{X(t)\}$ is recurrent if and only if the Markov chain $\{\theta_n\}$ is recurrent.*
4. *The PDMP $\{X(t)\}$ is positive recurrent if and only if the Markov chain $\{\theta_n\}$ is recurrent with invariant measure π satisfying $\pi H(M) < \infty$.*
5. *The PDMP $\{X(t)\}$ is Harris recurrent if and only if the Markov chain $\{\theta_n\}$ is Harris recurrent.*

The authors in [CD08] also give sufficient conditions on G (in a modified Foster-Lyapunov criterion form) to ensure the existence of an invariant probability measure, positive Harris recurrence and ergodicity for the PDMP.

In [Las04], stability and ergodicity via Meyn-Tweedie arguments are established for one dimensional PDMPs as AIMD (Additive Increase Multiplicative Decrease) processes. Most of these results are essentially of qualitative type, i.e. no practical information can be obtained

on the rate of convergence to equilibrium. The papers [CMP10, FGM12] are first attempts to get quantitative results for the long time behavior of special PDMPs: TCP (which is defined in Section 4) for the first one and a PDMP describing the motion of a bacteria for the second one.

3 Regularity and general convergence results for Markov switching systems

3.1 The Markov switching model

We consider a subclass of PDMP sometimes known as "Markov switching model". The "natural" state space of this process is the product space $\mathbb{R}^d \times \{1, \dots, n\}$ — as was said in Remark 1, this fits within the general theory described above. To define the process, consider n vector fields F^1, F^2, …, F^n on $\mathbb{R}^d$. These fields define flows ϕ^i via the ODE:

$$\frac{dy_t}{dt} = F^i(y_t)dt.$$

To simplify matters we suppose that the F^i are $\mathcal{C}^1$ and that there is a compact set $K \subset \mathbb{R}^d$ that is left invariant by all flows (nasty things may occur if this is not the case, see e.g. [BLBMZ12a]).

We also suppose that we are given n^2 nonnegative functions $\lambda_{ij} : \mathbb{R}^d \to \mathbb{R}$ (the jump rates), such that $\lambda_{ii}(y) = 0$, and for any given y, $(\lambda_{ij}(y))_{i,j}$ is irreducible.

The Markov switching model is a process $X_t = (Y_t, I_t) \in \mathbb{R}^d \times \{1, 2, \cdots n\}$ defined informally as follows:

- Y_t is driven by the vector field F^{I_t},
- if $I_t = i$ and $Y_t = y$, I_t jumps to j with rate $\lambda_{i,j}(y)$.

The corresponding generator is:

$$Lf(y,i) = F^i \cdot \nabla_x f + \sum_j \lambda_{ij}(y)(f(y,j) - f(y,i)).$$

This process can be constructed as described in Section 2. Another possible way is to generate a Poisson process with a high enough intensity to "propose" jump times, and to accept or reject these jumps with probabilities depending only on the values of $\lambda_{ij}(y)$ at the current point (see Section 2 of [BLBMZ12b] for details). This naturally defines a discrete process $\tilde{X}_n$ (a Markov chain), which turns out to be an alternative to the chain θ_n defined above, in the sense that results similar to Theorem 1 hold.

In this model, the trajectory of the "position" $Y_t \in \mathbb{R}^d$ does not jump, it evolves continuously. The jumps only occur on the discrete part I_t, the one that dictates which (deterministic) dynamics the position must follow. Let us also note that there are no "boundary jumps" in this model: for every $x = (y, i)$, the quantity $t^\star(x)$ defined by (1) is infinite.

Many questions that seem intractable for fully general PDMPs can be answered in this restricted framework. In particular, a kind of regularity for the law of (Y_t, I_t) can be established if the vector fields satisfy a condition that closely resembles Hörmander's classical condition for diffusion. These results are described in the following section. In the next one we discuss results on the "infinite time" behaviour of the process: once more, exponential convergence to equilibrium can be proved under fairly weak assumptions.

3.2 Regularity results

We are interested here in criteria that guarantee a regularity for the law of the continuous part, Y_t.

For the sake of comparison, let $B_t = (B_t^1, \dots, B_t^n)$ be an n-dimensional Brownian motion, and consider the solution of the following (Stratonovich) SDE:

$$dY_t = F^0(Y_t)dt + \sum_{i=1}^{n} F^i(Y_t) \circ dB_t^i.$$

Depending on the F^i, the law of Y_t may or may not have a density with respect to Lebesgue measure. For example, if the F^i are constant, the law is regular if and only if the F^i span $\mathbb{R}^d$. For non-constant fields there is a well-known criterion for regularity, known as Hörmander's criterion, and expressed in terms of the F^i and their iterated Lie brackets. Thanks to this criterion, one may prove that the law of a kinetic process (Y, V) driven only by a Brownian motion on its velocity is immediately absolutely continuous, with a regular density (see *e.g.* [Vil09]).

In our case, there is no hope for such a strong result. Indeed, the probability that (Y_t, I_t) has not jumped at all between 0 and t is strictly positive, therefore the law of Y_t has a singular part – we will see a similar phenomenon in the study of the TCP process with the lower bound (5).

However, let us define the following families of vector fields:

$$\mathcal{V}_0 = \{F^i - F^j, i \neq j\},$$
$$\forall i \geq 0, \qquad \mathcal{V}_{i+1} = \mathcal{V}_i \cup \{[F^i, G], G \in \mathcal{V}_i\}.$$

With these notations the following result holds:

Theorem 2 (Absolute continuity, [BH12, BLBMZ12b])**.** *Suppose there is a point $p \in \mathbb{R}^d$ and an integer k such that the "bracket condition" holds, i.e. the iterated Lie brackets $\{V(p), V \in \mathcal{V}_k\}$ span $\mathbb{R}^d$. Then there is a time $t > 0$, a strictly positive η, and two probability measures ν_t^c, ν_t^r on $\mathbb{R}^d \times \{1, \dots, n\}$ such that:*

- *the law μ_t of Y_t starting from p can be written as $\mu_t = \eta\nu_t^c + (1-\eta)\nu_t^r$,*
- *the "continuous" part ν_t^c is absolutely continuous with respect to the product of the Lebesgue measure on $\mathbb{R}^d$ and the counting measure on $\{1, \dots, n\}$.*

Similar results are obtained in [BH12, BLBMZ12b] for the discrete chain $\tilde{Z}_n$; they involve another (weaker) bracket condition.

3.3 Invariant measures and convergence

To study the long time behaviour of (Y_t, I_t), it is natural to look for an analogue of the ω-limit set (defined when there is only one flow). To that end, first define the positive trajectory of y as the set

$$\gamma^+(y) = \bigcup_{m\in\mathbb{N}} \left\{ \phi^{i_m}_{t_m}(\cdots(\phi^{i_1}t_1(y))\cdot) \middle| (i_1,\ldots,i_m) \in \{1,\ldots,n\}^m; (t_1,\ldots,t_m) \in \mathbb{R}_+^m \right\}.$$

That is, $\gamma^+(y)$ is the set of points reachable from y by a trajectory of the process. Recall that all the F^i leave a compact set K invariant. The right limit set to consider is the *accessible set*, the (possibly empty) compact set $\Gamma \subset K$ defined as

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

Informally this is the set of points that are (approximately) reachable from anywhere in K by following the flows.

This set is deeply connected with the invariant measures of the process. Let $\mathcal{P}_{inv}$ be the set of probability measures that are invariant for the process, and let $E = \{1,\ldots,n\}$ be the finite index set.

Proposition 1 (Limit set and support of invariant measures).

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

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

(iii) Suppose that $\mathcal{P}_{inv}$ is a singleton $\{\pi\}$. Then $supp(\pi) = \Gamma \times E$.

Once more, similar statements hold for the invariant measures of the discrete chain $\tilde{Z}_n$.

A usual way of proving convergence to equilibrium for Markov chains or processes, following the classical idea of the Foster-Lyapunov criterion, is to identify a "good" subset A of the state space, satisfying two properties:

- the process has a tendency to return to A;
- for any x, y in A, the laws of the process starting from x and y are similar in some sense.

If these two properties hold, one can construct a coupling between trajectories in three steps:

- wait for the two processes to come back to A;
- once they are in A, using the second property to try to couple them;
- if the coupling fails, go back to step 1.

In the classical case of a discrete Markov chain, A will usually be a single point so that the second property will be trivial; the "tendency to return" will typically follow from the existence of a Lyapunov function. This type of technique is refined and generalized for Markov chains on general state spaces in the renowned book by Meyn and Tweedie [MT93].

In our case, without additional assumptions, we cannot hope to find a Lyapunov function. However, if Γ is not empty, and if p is a point in Γ, we can prove — using compactness arguments — that the process will visit infinitely often any neighborhood of p.

To use the heuristics above, it remains to show that two copies of the process, starting from a point near p, have a good chance to couple. This is where the regularity results may be used: two variables with strictly positive density can be coupled with positive probability. This is the main idea of the proof of:

Theorem 3 (Exponential convergence, [BLBMZ12b])**.** *Suppose that the "bracket condition" of Theorem 2 is satisfied at a point $p \in \Gamma$. Then there exists a unique invariant probability measure μ, its support is $\Gamma \times \{1, \dots, n\}$, and $X_t = (Y_t, I_t)$ converges exponentially fast in total variation norm:*

$$\exists\, \alpha > 0, C > 0, \quad \|\mathcal{L}(X_t) - \pi\|_{\mathrm{TV}} \leq C \exp(-\alpha t). \tag{2}$$

These convergence results are quite general. However, there is no hope to find reasonable convergence rates by this approach, since we used compactness arguments. In particular cases, one can often use the specific structure of the process to prove much better bounds. This is done in the Markov switching case under additional assumptions in [BLBMZ12c]. The next section describes similar results for the slightly different case of the TCP window-size process.

4 Explicit convergence rates via coupling

Beyond the problems of existence, uniqueness, and regularity of an invariant probability measure, an important and challenging question is to get explicit and efficient rates of convergence to equilibrium for PDMPs. It seems hard to address the problem via spectral methods since a PDMP is generally, and inherently, non reversible (see however [GM13, MM13] for positive results in this direction). As mentioned previously, coupling methods "à la" Meyn-Tweedie can be applied but hardly give efficient rates of convergence. We will present here a specific method that allows to treat one specific (and relatively simple) example of PDMP, the so-called TCP window-size process.

4.1 The TCP window-size process

The TCP window-size process appears as a scaling limit of the transmission rate of a server uploading packets on the Internet according to the algorithm used in the TCP (Transmission Control Protocol) in order to avoid congestions (see [DGR02] for details on this scaling limit).

This process $X = \{X_t\}_{t\geq 0}$ has $[0,\infty)$ as state space and its infinitesimal generator is given, for any smooth function $f : [0,\infty) \to \mathbb{R}$, by

$$Lf(x) = f'(x) + x(f(x/2) - f(x)). \tag{3}$$

This means that the transmission rate is increased linearly when there is no congestion, and divided by two in case of congestion. A congestion occurs at random times, with rate given by the value of the transmission rate. This process is rather simple and its stationary regime is well understood: it is ergodic and admits a unique invariant probability measure μ, with an explicit density (see [DGR02]). Getting an explicit speed of convergence is however not so simple, since the TCP process is irreversible but also because, in some sense, it "lacks randomness" (the only randomness comes from the jump times), what makes any coupling procedure delicate, particularly for convergence in total variation norm. In the following, the semigroup associated to the TCP process will be denoted $\{P_t\}_{t\geq 0}$.

4.2 Speed of convergence in Wasserstein distance

The speed of convergence in Wasserstein distance can be dealt with using a coupling first introduced in [CMP10]. This is a Markov process on $\mathbb{R}_+^2$ whose marginals are two TCP processes, and which is defined by its generator as follows: if $x \geq y$,

$$\mathfrak{L}f(x,y) = (\partial_x + \partial_y)f(x,y) + y\Big(f(x/2,y/2) - f(x,y)\Big) + (x-y)\Big(f(x/2,y) - f(x,y)\Big), \tag{4}$$

and it is given by the symmetric expression if $x > y$. In words, with a rate equal to the minimum of the two coordinates, both trajectories jump simultaneously, and with a rate equal to the difference between the two coordinates the higher trajectory jumps alone (and the lower one never jumps alone). It induces problems since the higher trajectory may jump alone even when the distance between both trajectories is small (hence at a small but positive rate), increasing dramatically this distance. This can also be seen on computation, noting that this dynamics does not contract the distance between trajectories in L^p distance, for any $p \geq 1$. However, one gets (non uniform) contraction for $p = 1/2$. Working this out, and using also a bound on any moment of the process at any time uniformly in the starting point, one finally gets the following result:

Theorem 4 (Explicit rate for TCP in Wasserstein distance, [BCG$^+$13])**.** *Let $c = \sqrt{2}(3+\sqrt{3})/8 \approx 0.84$ and $\lambda = \sqrt{2}(1-\sqrt{c}) \approx 0.12$. For any $\tilde{\lambda} < \lambda$, any $p \geq 1$ and any $t_0 > 0$, there is a constant $C = C(p,\tilde{\lambda},t_0)$ such that, for any initial probability measures ν and $\tilde{\nu}$ and any $t \geq t_0$,*

$$W_p(\nu P_t, \tilde{\nu} P_t) \leq C \exp\left(-\frac{\tilde{\lambda}}{p} t\right).$$

This result can be compared to numerical simulations: the exponential rate of convergence for the W_1 Wasserstein distance to the invariant measure is about 1.6, and the one for the L^1 distance between both coordinates of the coupling we used is about 0.5 (which means that one cannot do better using this coupling).

4.3 Speed of convergence in total variation norm

Let us first note that, even starting from two arbitrarily close initial conditions, there is a positive probability that two TCP processes never have jumped before a given time t, which gives an immediate and striking lower bound for the total variation distance between laws at time t. To be more specific, the first jump time T_1 of the TCP process starting from x is characterized by $p_t(x) = \mathbb{P}\{T_1 \geq t\} = e^{-t^2/2-xt}$. This implies that

$$\|\delta_x P_t - \delta_y P_t\|_{\mathrm{TV}} \geq p_t(x) \vee p_t(y) = e^{-t^2/2-(x\wedge y)t}\,. \tag{5}$$

Consequently, this distance is not tending to 0 as y tends to x. Nevertheless, it can be shown that, for y close to x, the distance is really on the order of this lower bound. To note that, we construct a coupling which tries to stick two TCP processes in one jump. Indeed, assume that $x > y$ and denote by T_1^x the first jump time of the process starting from x and T_1^y, T_2^y the two first jump times of a TCP process starting from y. Then if these random variables satisfy the two relations

$$T_1^x = T_1^y + x - y \tag{6}$$

and $T_2^y - T_1^y \geq x - y$, both processes are at the same position at time T_1^x, hence can be coupled from this time. The higher probability of getting (6) is an optimal coupling probability for two measures with density, it can hence be computed explicitly, and gives an upper bound on the order of (5).

We finally can construct a coupling in the following way: for any $\varepsilon > 0$, we use the dynamical coupling defined in (4) during a time t_1 on the order of $\log(1/\varepsilon)$ to try to make the trajectories ε close; we then try to stick them in a time t_2 on the order of $\sqrt{\log(1/\varepsilon)}$. Taking ε to 0 gives the following result:

Theorem 5 (Explicit rate for TCP in total variation, [BCG$^+$13])**.** *Let λ as defined in Theorem 4. For any $\tilde{\lambda} < \lambda$ and any $t_0 > 0$, there exists C such that, for any initial probability measures ν and $\tilde{\nu}$ and any $t \geq t_0$,*

$$\|\nu P_t - \tilde{\nu} P_t\|_{\mathrm{TV}} \leq C \exp\left(-\frac{2\tilde{\lambda}}{3} t\right).$$

4.4 Application to other processes

This strategy (first apply a coupling which is efficient for the Wasserstein distance, then stick the trajectories in one single attempt) can also be used for other processes. We can treat for example the case of the TCP process with constant jump rate, with generator

$$Lf(x) = f'(x) + r(f(x/2) - f(x)), \tag{7}$$

$r > 0$ being fixed. This gives a purely probabilistic proof (and slightly better constants) of results first obtained in an analytic way by Perthame and Ryzhik in [PR05].

The same idea can also be applied to diffusion processes, associating recent results from [Ebe11, BGG12] on the convergence in Wasserstein distance with results on the regularity of the semigroup at small time from [Wan10]. The reader is referred to [BCG$^+$13] for details.

5 Infinite dimensional PDMP

5.1 Spatially extended conductance based neuron model

The intrinsic variability of a large class of biological systems is the subject of an intensive area of research since a few years. Spatially extended conductance based models of biological excitable membranes, and more particularly of neuronal membranes [HH52], are considered, at an appropriate scale, as hybrid stochastic models [Aus08, FWL05]. More precisely, these models can be regarded as belonging to the subclass of stochastic hybrid models consisting of PDMPs [BR11].

The neurons are the basic building blocks of the nervous systems. They are, roughly speaking, made of three parts: the dendrites, the soma and the axon. In the present discussion, we focus our attention on the axon, also referred as nerve fiber. The main role of the axon is to transmit, for example to another neuron, the signal it received from the soma. This electrical signal is called the nerve impulse or the action potential or the spike. The target neuron may be far from the neuron which emits the spike, therefore the nerve fiber is often much longer than width and for mathematical convenience we model the axon by a segment, here the segment $I = [0, 1]$. As for all biological cells, the neuronal membrane allows the exchanges of ions between the internal and the external cellular media through the ionic channels. We assume that these ionic channels are disposed at discrete location along the axon: at a location $z_i \in \mathbb{Z}$ is a ionic channel where $\mathbb{Z}$ is a finite subset of $(0, 1)$ of cardinality N. When an ionic channel is open, it allows a flow of ions to enter or leave the cell, that is, it allows a current to pass. We denote by $\xi \in E$, where E is a finite state space, the state of the channel at location z_i. Basically a state is to be open or closed. When open, an ionic channel in the state ξ allows the ions of a certain specie associated to ξ to pass through the membrane. If we write $u(z_i)$ for the local potential of the neuronal membrane, often called the trans-membrane potential, then the current which passes through the membrane is given by

$$c_\xi(v_\xi - u(z_i))$$

where c_ξ is the conductance of the ionic channel in state ξ and v_ξ the associated driven potential. The quantities c_ξ and v_ξ are both real constants, the first being positive. The sign of $v_\xi - u(z_i)$ tells us if the ions corresponding to the state ξ are leaving or entering the cell. The rate at which the channel at location z_i goes from one state ξ to another state ζ depends on the local trans-membrane potential $u(z_i)$ and is denoted in the sequel by $q_{\xi\zeta}(u(z_i))$. For example, when the local potential $u(z_i)$ is high, the ionic channels in a specific state ξ will have the tendency to open whereas when $u(z_i)$ is low, the ionic channels in the same state ξ will have the tendency to close. This is this mechanism which allows the creation of an action potential and gradually its propagation along the axon. Let C be the capacitance of the axon membrane, a the radius of the axon and R its internal resistance, this three quantities being positive constants. The classical conductance based model for the propagation of the

nerve impulse is the cable equation

$$C\partial_t u = \frac{a}{2R}\Delta u + \frac{1}{N}\sum_{i\in\mathbb{Z}} c_{r(i)}(v_{r(i)} - u(z_i))\delta_{z_i} + \mathcal{I} \tag{8}$$

where $r(i)$ denotes the state of the channel at location z_i. The dynamic of the channel $r(i)$ is given by

$$\mathbb{P}\left[r_{t+h}(i) = \zeta | r_t(i) = \xi\right] = q_{\xi\zeta}(u_t(z_i))h + \mathrm{o}(h) \tag{9}$$

where the ionic channels $\{r(i), i \in \mathbb{Z}\}$ evolve independently over infinitesimal time-scales. The quantity $\mathcal{I}$ in (8) corresponds to the impulse received by the axon. Since it does not affect our analysis, we state $\mathcal{I} = 0$ in the sequel. At last, the Partial Differential Equation (8) has to be endowed with boundary conditions and we consider in the present paper zero Dirichlet boundary conditions.

Let us be more specific about the assumptions on the jump rates of the ionic channels. We assume that the function $q_{\xi\zeta}$ are smooth and take values between two fixed positive constants. This assumption certainly holds in the classical models for excitable membrane such as the Hodgkin-Huxley model, Morris-Lecar model or models in Calcium dynamics. To gain a more accurate insight in conductance based neuron model we refer to the classical book [Hil84].

Let us denote by $r = (r(i), i \in \mathbb{Z})$ the configuration of the whole ionic channels: it is a jump process with values in E^N. For a given fixed trans-membrane potential u, r follows the following dynamic. For $r \in E^N$, the total rate of leaving state r is given by

$$\Lambda(u, r) = \sum_{i\in\mathbb{Z}} \sum_{\xi \neq r(i)} q_{r(i),\xi}(u(z_i))$$

and, if another state r' differs from r only by the component $r(i_0)$, then r' is reached with probability

$$Q(u, r)(\{r'\}) = \frac{q_{r(i_0),r'(i_0)}(u(z_{i_0}))}{\Lambda(u, r)}$$

If r' differs from r by two or more components then $Q(u, r)(\{r'\}) = 0$. Between each jump of the process r, the potential u follows the deterministic dynamic induced by the PDE (8). For a fixed time horizon T, the process $((u_t, r_t), t \in [0, T])$ described this way is a piecewise deterministic process (PDP) with values in the infinite dimensional space[1] $H_0^1(I) \times E^N$. This PDP is Markov — and therefore a PDMP — if one chooses an appropriate filtration:

Theorem 6 ([BR11])**.** *There exists a filtered probability space satisfying the usual conditions such that the process (u_t, r_t) is a homogeneous Markov process on $H_0^1(I) \times E^N$,*

[1] $H_0^1(I)$ denotes the usual Sobolev space of functions in $L^2(I)$ with first derivative in the sens of the distributions also in $L^2(I)$ and with trace equals to zero on the boundary $\{0, 1\}$ of I.

5.2 Singular perturbation of conductance based neuron model

In this section we consider the spatially extended conductance based model of the previous section as a slow-fast system: some states of the ion channels communicate faster between them than others. This phenomenon is intrinsic to classical neuron model, see [Hil84]. Mathematically, this leads to introduce a parameter $\epsilon > 0$, destined to be small, in equations (8) and (9). For the states which communicate at a faster rate, we say that they communicate at the usual rate divided by ϵ.

The state space E is therefore partitioned into different classes of states which cluster the states communicating at a high rate. This kind of description is usual, see for example [FGC10]. We regroup our states in classes making a partition of the state space E into

$$E = E_1 \sqcup \cdots \sqcup E_l$$

where $l \in \{1, 2, \cdots\}$ is the number of classes. Inside a class E_j, the states communicate faster at jump rates of order $\frac{1}{\epsilon}$. States of different classes communicate at the usual rate of order 1. For $\epsilon > 0$ fixed, we denote by (u^ϵ, r^ϵ) the modification of the PDMP introduced in the previous section with now two time scales. The new equations for the two time-scales model are

$$C\partial_t u^\epsilon = \frac{a}{2R}\Delta u^\epsilon + \frac{1}{N}\sum_{i\in\mathbb{Z}} c_{r^\epsilon(i)}(v_{r^\epsilon(i)} - u^\epsilon(z_i))\delta_{z_i} \tag{10}$$

that is there is no changes in the form of the equation on the potential whereas the dynamic of the channel $r^\epsilon(i)$ is given by

$$\mathbb{P}\left[r^\epsilon_{t+h}(i) = \zeta | r^\epsilon_t(i) = \xi\right] = q^\epsilon_{\xi\zeta}(u^\epsilon_t(z_i))h + \mathrm{o}(h) \tag{11}$$

where $q^\epsilon_{\xi\zeta} = \frac{1}{\epsilon}q_{\xi\zeta}$ if ξ and ζ are in the same class and $q^\epsilon_{\xi\zeta} = q_{\xi\zeta}$ otherwise.

For any fixed potential $u(z_i) = y$ we assume that the jump process $r(i)$ restricted to the class E_j is irreducible and has therefore a unique quasi-stationary distribution $\mu_j(y)$. The measure $\mu_j(y)$ does not depend on i because when the potential is held fixed to y, all the channels $\{r(i), i \in \mathbb{Z}\}$ have same law. The main idea is that when ϵ goes to zero, the fast components of the system (10)-(11) reach their stationary behavior that can be viewed as an averaged dynamic. For the jump process r with a potential held fixed to y, this leads to consider the averaged jump process $\bar{r} = (\bar{r}(i), i \in \mathbb{Z})$ with values in $\{1, \cdots, l\}^N$ and jump rates between two different classes given by

$$\bar{q}_{l_1 l_2}(y) = \sum_{\zeta\in E_{l_1}}\sum_{\xi\in E_{l_2}} q_{\zeta\xi}(y)\mu_{l_1}(y)(\zeta)$$

$\bar{q}_{l_1 l_2}(y)$ is indeed the sum of the rate of jumps from one state ζ in E_{l_1} to a state ξ in E_{l_2} averaged against the quasi-stationary measure associated to the class E_{l_1}. The same idea can be applied to the equation on the potential (10) and furthermore we have the following result.

Theorem 7 ([GT12]). *The process $(u_t^\epsilon, t \in [0,T])$ converges in law in $\mathcal{C}([0,T], H_0^1(I))$ when ϵ goes to zero toward the process u solution of*

$$C\partial_t u = \frac{a}{2R}\Delta u + \bar{G}_{\bar{r}}(u) \tag{12}$$

with zero Dirichlet boundary conditions. The averaged reaction term is given by

$$\frac{1}{N}\sum_{i\in\mathbb{Z}}\sum_{j=1}^{l} 1_j(\bar{r}(i)) \sum_{\xi\in E_j} \mu_j(u(z_i))(\xi)c_\xi(v_\xi - u(z_i))\delta_{z_i}$$

and the dynamic of the averaged ionic channels is

$$\mathbb{P}\left[\bar{r}_{t+h}(i) = l_2 | \bar{r}_t(i) = l_1\right] = \bar{q}_{l_1 l_2}(u_t(z_i))h + \mathrm{o}(h) \tag{13}$$

The above theorem can be seen as a law of large number as ϵ tends to zero. To continue our analysis, we study the fluctuations around the averaged limit with a Central Limit Theorem (CLT).

Theorem 8. *The fluctuations $z^\epsilon = \frac{u-u^\epsilon}{\sqrt{\epsilon}}$ converge in law in $\mathcal{C}([0,T], H_0^1(I))$ when ϵ goes to zero toward a process z corresponding to the solution of a Hybrid Stochastic Partial Differential Equation (HSPDE).*

One of the main advantages of considering the averaged model (12)-(13) instead of the two time-scales model (10)-(11) is that the dynamic of the ionic channels is simplified: the state space for the averaged ionic channels has a smaller cardinality than the original state space. Averaging reduces in this case the dimension of the model. This can be very useful for simulations or bifurcation analysis. However, we notice that the averaged PDE on the potential (12) may appear to be much more non linear than in the non averaged case (10). We also remark that the error made by averaging is controlled by the CLT and that an associated Langevin equation can be derived. The Langevin approximation consists also in a HSPDE. If well posed and tractable, the Langevin approximation may be regarded as a good compromise between the two-time scales and the averaged model.

6 Estimation of the distribution of the inter-jumping times

This section is dedicated to nonparametric statistics for piecewise deterministic Markov processes. A suitable choice of the state space and the main characteristics of the process provide stochastic models covering a large number of applications, in reliability for instance or in biology as mentioned before. In this context, it seems essential to develop estimation methods for this class of processes.

In this part, we consider a piecewise deterministic Markov process $\{X(t)\}$ defined on an open subset $\mathcal{X}$ of a separable metric space $(\mathcal{E}, d)$. The conditional distribution of the inter-jumping time S_{i+1} given the previous post-jump location Z_i is given by

$$\forall t \geq 0, \quad \mathbb{P}\left[S_{i+1} > t | Z_i\right] = \exp\left(-\int_0^t \lambda(\phi(Z_i, s))\, ds\right) \mathbf{1}_{\{0 \leq t < t^\star(Z_i)\}},$$

where we recall that $t^\star(x)$ is the deterministic exit time from $\mathcal{X}$ starting from x. One may associate with the compound function $\lambda \circ \phi$ the conditional density f of the inter-jumping times, given for any $(x, t) \in \mathcal{X} \times \mathbb{R}_+$ by:

$$f(x, t) = \lambda(\phi(x, t)) \exp\left(-\int_0^t \lambda(\phi(x, s))\, ds\right).$$

We investigate here the nonparametric estimation of $f(x, t)$ for any $x \in \mathcal{X}$ and $0 \leq t < t^\star(x)$ from the observation of one trajectory of the process within a long time. In the context of piecewise deterministic Markov processes, all the randomness is contained in the embedded Markov chain (Z_n, S_n). As a consequence, our statistics may be computed from the post-jump locations Z_n and the inter-jumping times S_n. A precise formulation of the results and the investigated methods may be found in [ADGP12a]. The general framework of metric state space has been chosen in order to provide an estimate which could be computed in applications involving state spaces of the form $\mathcal{X} = E \times M$, where M is continuous and E is discrete.

This study is premised on a previous work of Azaïs and al. In [ADGP12b], the authors investigate the nonparametric estimation of the jump rate for a class of nonhomogeneous marked renewal processes. This kind of stochastic models may be directly related with piecewise deterministic Markov processes with constant flow. For this class of processes, one may exhibit some continuous-time martingales, which appear in the well-known framework of Aalen's multiplicative intensity model (see [Aal75, Aal77, Aal78]). Nevertheless, this idea is not relevant for studying the estimation of the jump rate for piecewise deterministic Markov processes, since the conditional independence $\mathbb{E}\left[g(Z_{n+1}) | Z_n\right] = \mathbb{E}\left[g(Z_{n+1}) | Z_n, S_{n+1}\right]$ is not satisfied in the general case. Consequently, we propose to examine the conditional distribution of S_{n+1} given the previous post-jump location Z_n and the next post-jump location Z_{n+1}. We prove that there exists, under some regularity assumptions, a mapping $\widetilde{\lambda}$ from $\mathcal{X} \times \mathcal{X} \times \mathbb{R}_+$ into $\mathbb{R}_+$, such that for any $t \geq 0$,

$$\mathbb{P}\left[S_{n+1} > t | Z_n, Z_{n+1}\right] = \exp\left(-\int_0^t \widetilde{\lambda}(Z_n, Z_{n+1}, s)\, ds\right) \mathbf{1}_{\{0 \leq t < t^\star(Z_n)\}}.$$

In addition, we state that there exists a structure of continuous-time martingale, which is similar to the one of the multiplicative intensity model. Thanks to these considerations, we provide a way to estimate an approximation l of the jump rate $\widetilde{\lambda}$.

Our strategy for estimating the conditional probability density function $f(x, t)$ consists in the introduction of a partition (B_k) of the state space and two functions $l(A, B_k, t)$ and

$H(A, B_k, t)$, where A is a set containing x. (The set A is not allowed to intersect the boundary of the state space $\mathcal{X}$.) They provide a way to approximate the function of interest. Indeed, $f(x,t)$ is close to

$$\sum_k l(A, B_k, t) H(A, B_k, t),$$

if the partition (B_k) and the set A are thin enough. One may give an interpretation of these two functions. First, $l(A, B_k, t)$ is an approximation of the jump rate $\widetilde{\lambda}(x, y, t)$ of the inter-jumping time S_{n+1} given $Z_n = x$ and $Z_{n+1} = y$, under the stationary regime, for $x \in A$ and $y \in B_k$. Roughly speaking, $l(A, B_k, t)$ may be seen as the jump rate from A to B_k at time t under the stationary regime. Furthermore, $H(A, B_k, t)$ is exactly the conditional probability $\mathbb{P}_\nu [S_{n+1} > t, Z_{n+1} \in B_k | Z_n \in A]$, where ν denotes the invariant measure of the Markov chain (Z_n). We propose to estimate both the functions $l(A, B_k, t)$ and $H(A, B_k, t)$ in order to provide a consistent estimator of the density of interest $f(x,t)$.

Firstly, our work is inspired by the smoothing methods proposed by Ramlau-Hansen in [RH83] and the previous work of Azaïs et al. [ADGP12b]. We provide a nonparametric kernel estimator $\widehat{l}_n(A, B_k, t)$ of $l(A, B_k, t)$. The keystone to state the consistency of this estimator lies in the structure of continuous-time martingale mentioned above. Furthermore, we estimate the conditional probability $H(A, B_k, t)$ by its empirical version

$$\widehat{H}_n(A, B_k, t) = \frac{\sum_{i=0}^{n-1} \mathbf{1}_{\{Z_{i+1} \in B_k\}} \mathbf{1}_{\{S_{i+1} > t\}} \mathbf{1}_{\{Z_i \in A\}}}{\sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}}}.$$

The error between the function of interest and its estimate breaks down into two parts: a deterministic error caused by the approximation of $f(x,t)$ and a stochastic one in the estimation of this approximation. Our major result of consistency is stated in the following proposition.

Proposition 2. For any $\varepsilon, \eta > 0$, there exist an integer N, a set A and a partition (B_k) such that, for any $n \geq N$ and $0 < r_1 < r_2 < \inf_{x \in A} t^\star(x)$,

$$\mathbb{P}\left[\sup_{r_1 \leq t \leq r_2} \left| f(x,t) - \sum_k \widehat{l}_n(A, B_k, t) \widehat{H}_n(A, B_k, t) \right| > \eta \right] < \varepsilon.$$

In addition, one may state the uniform convergence on every compact subset $K \subset \mathcal{X}$. This nonparametric estimator of the conditional distribution of the inter-jumping times is easy to compute from the observation of the embedded chain (Z_n, S_n) within a long time. However, the choice of the set A and the partition (B_k) remains an open issue.

7 Numerical methods.

A lot of numerical methods have been developed recently to simulate diffusion processes and compute expectations, stopping times and other interesting quantities (see references in [dSDG10] for example). But since PDMPs are in essence discontinuous at random times,

these results turn out to be too specific to be applied to them. Besides, another important source of complication is the fact that the transition semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ of $\{X(t)\}$ cannot be explicitly computed from the local characteristics (ϕ, λ, Q) of the PDMP (see [DC99, CD08]). Therefore, it turns out to be hard to give an explicit expression for the Markov kernel P associated with the Markov chain $\{X(t)\}$. Also, the Markov chain $\{X(t)\}$ is, in general, not even a Feller chain (see [Dav93], pages 76 and 77).

On the other hand, PDMPs exhibit some nice specific properties. For instance, all the randomness of the process can be described by the discrete time (continuous state space) Markov chain $\{\Theta_n\}$ of the post jump locations and inter-jump times.

There exist essentially two families of numerical methods. The first one is based on the discretization the Markov kernel Q [CD88, CD89, Cos93]. The second one relies on a discretization of the Markov chain $\{\Theta_n\}$ using quantization [Pag98].

Thanks to these methods one can solve optimal stopping problems [CD88, dSDG10, Gug86], impulse control problems [Cos93, CD89, dSD12], approximate distributions of exit times [BDSD12b] and compute expectations of functionals of PDMP's [BdSD12a].

Acknowledgements. The authors warmly thank Benoîte de Saporta and Florent Malrieu for their kind and useful advices on the redaction of these proceedings. They also thank Arnaud Guillin for his enthusiastic and stimulating organization of the "Journées MAS 2012" in Clermont-Ferrand and of their subsequent proceedings. The research of J.-B Bardet, A. Génadot, N. Krell and P.-A. Zitt is partly supported by the Agence Nationale de la Recherche PIECE 12-JS01-0006-01. The research of Romain Azaïs was supported by ARPEGE program of the Agence Nationale de la Recherche, project FAUTOCOES, number ANR-09-SEGI-004.

References

[Aal75] O. O. Aalen, *Statistical inference for a family of counting processes*, ProQuest LLC, Ann Arbor, MI, 1975, Thesis (Ph.D.)–University of California, Berkeley. MR 2625917

[Aal77] ———, *Weak convergence of stochastic integrals related to counting processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **38** (1977), no. 4, 261–277. MR 0448552 (56 #6858)

[Aal78] ———, *Nonparametric inference for a family of counting processes*, Ann. Statist. **6** (1978), no. 4, 701–726. MR 491547 (80a:62137)

[ADGP12a] R. Azaïs, F. Dufour, and A. Gégout-Petit, *Nonparametric estimation of the conditional distribution of the inter-jumping times for piecewise-deterministic Markov processes*, Preprint arXiv:1202.2212v2, 2012.

[ADGP12b] ———, *Nonparametric estimation of the jump rate for non-homogeneous marked renewal processes*, Preprint arXiv:1202.2211v2, Accepted for publication in Annales de l'Institut Henri Poincaré, 2012.

[Aus08] T. D. Austin, *The emergence of the deterministic Hodgkin-Huxley equations as a limit from the underlying stochastic ion-channel mechanism*, Ann. Appl. Probab. **18** (2008), no. 4, 1279–1325. MR 2434172 (2010a:60113)

[BCG+13] J.-B. Bardet, A. Christen, A. Guillin, A. Malrieu, and P.-A. Zitt, *Total variation estimates for the TCP process*, Electron. J. Probab. **18** (2013), no. 10, 1–21.

[BdSD12a] A. Brandejsky, B. de Saporta, and F. Dufour, *Numerical method for expectations of piecewise deterministic markov processes*, CAMCoS **7** (2012), 63–104.

[BDSD12b] A. Brandejsky, B. De Saporta, and F. Dufour, *Numerical methods for the exit time of a piecewise-deterministic Markov process*, Adv. in Appl. Probab. **44** (2012), no. 1, 196–225. MR 2951552

[BGG12] F. Bolley, I. Gentil, and A. Guillin, *Convergence to equilibrium in Wasserstein distance for Fokker–Planck equations*, J. Funct. Anal. **263** (2012), no. 8, 2430–2457. MR 2964689

[BH12] Y. Bakhtin and T. Hurth, *Invariant densities for dynamical systems with random switching*, Nonlinearity **25** (2012), no. 10, 2937–2952.

[BLBMZ12a] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt, *On the stability of planar randomly switched systems*, to appear in Ann. Appl. Probab. Preprint arXiv:1204.1921 available on arXiv, 2012.

[BLBMZ12b] ———, *Qualitative properties of certain piecewise deterministic Markov processes*, Preprint arXiv:1204.4143 available on arXiv, 2012.

[BLBMZ12c] ———, *Quantitative ergodicity for some switched dynamical systems*, Electron. Commun. Probab. **17** (2012), no. 56, 1–13.

[BR11] E. Buckwar and M. G. Riedler, *An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution*, J. Math. Biol. **63** (2011), no. 6, 1051–1093.

[CD88] O. L. V. Costa and M. H. A. Davis, *Approximations for optimal stopping of a piecewise-deterministic process*, Math. Control Signals Systems **1** (1988), no. 2, 123–146. MR 936330 (89h:90249)

[CD89] ———, *Impulse control of piecewise-deterministic processes*, Math. Control Signals Systems **2** (1989), no. 3, 187–206. MR 997213 (90i:93146)

[CD08] O. L. V. Costa and F. Dufour, *Stability and ergodicity of piecewise deterministic Markov processes*, SIAM J. Control Optim. **47** (2008), no. 2, 1053–1077. MR 2385873 (2009b:93163)

[CMP10] D. Chafaï, F. Malrieu, and K. Paroux, *On the long time behavior of the TCP window size process*, Stochastic Process. Appl. **120** (2010), no. 8, 1518–1534.

[Cos93] O. L. V. Costa, *Discretizations for the average impulse control of piecewise deterministic processes*, J. Appl. Probab. **30** (1993), no. 2, 405–420. MR 1212672 (94d:93143)

[Dav84] M. H. A. Davis, *Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models*, J. Roy. Statist. Soc. Ser. B **46** (1984), no. 3, 353–388, With discussion. MR MR790622 (87g:60062)

[Dav93] ———, *Markov models and optimization*, Monographs on Statistics and Applied Probability, vol. 49, Chapman & Hall, London, 1993.

[DC99] F. Dufour and O. L. V. Costa, *Stability of piecewise deterministic Markov processes*, SIAM J. Control Optim. **37** (1999), no. 5, 1483–1502 (electronic).

[DD02] F. Dufour and Y. Dutuit, *Dynamic reliability: A new model*, Proceedings of ESREL 2002 Lambda-Mu 13 Conference, 2002, pp. 350–353.

[DGR02] V. Dumas, F. Guillemin, and Ph. Robert, *A Markovian analysis of additive-increase multiplicative-decrease algorithms*, Adv. in Appl. Probab. **34** (2002), no. 1, 85–111.

[dSD12] B. de Saporta and F. Dufour, *Numerical method for impulse control of piecewise deterministic Markov processes*, Automatica J. IFAC **48** (2012), no. 5, 779–793. MR 2912800

[dSDG10] B. de Saporta, F. Dufour, and K. Gonzalez, *Numerical method for optimal stopping of piecewise deterministic Markov processes*, Ann. Appl. Probab. **20** (2010), no. 5, 1607–1637. MR 2724397 (2011g:60079)

[Ebe11] A. Eberle, *Reflection coupling and Wasserstein contractivity without convexity*, C. R. Math. Acad. Sci. Paris **349** (2011), no. 19-20, 1101–1104.

[EO05] R. Erban and H. G. Othmer, *From individual to collective behavior in bacterial chemotaxis*, SIAM J. Appl. Math. **65** (2004/05), no. 2, 361–391 (electronic). MR 2123062 (2005j:35220)

[FGC10] A. Faggionato, D. Gabrielli, and M. R. Crivellari, *Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors*, Markov Process. Related Fields **16** (2010), no. 3, 497–548. MR 2759771 (2011g:60047)

[FGM12] J. Fontbona, H. Guérin, and F. Malrieu, *Quantitative estimates for the long time behavior of an ergodic variant of the telegraph process*, Adv. in Appl. Probab. **44** (2012), no. 4, 977–994.

[FWL05] A. A. Faisal, J. A. White, and S. B. Laughlin, *Ion-channel noise places limits on the miniaturization of the brain'??s wiring*, Current Biology **15** (2005), no. 12, 1143–1149.

[GM13] S. Gadat and L. Miclo, *Spectral decompositions and $\mathbb{L}^2$-operator norms of toy hypocoercive semi-groups*, Kinet. Relat. Models **6** (2013), no. 2, 317–372. MR 3030715

[GRZ04] F. Guillemin, P. Robert, and B. Zwart, *AIMD algorithms and exponential functionals*, Ann. Appl. Probab. **14** (2004), no. 1, 90–117. MR 2023017 (2004m:60216)

[GT12] A. Genadot and M. Thieullen, *Averaging for a fully coupled piecewise-deterministic markov process in infinite dimensions*, Advances in Applied Probability **44** (2012), no. 3, 749–773.

[Gug86] U. S. Gugerli, *Optimal stopping of a piecewise-deterministic Markov process*, Stochastics **19** (1986), no. 4, 221–236. MR 872462 (88j:60084)

[HH52] A. L. Hodgkin and A. F. Huxley, *A quantitative description of membrane current and its application to conduction and excitation in nerve*, J. Physiol. **117** (1952), no. 4, 500.

[Hil84] B. Hille, *Ionic channels of excitable membranes*, vol. 174, Sinauer Associates Sunderland, MA, 1984.

[Las04] G. Last, *Ergodicity properties of stress release, repairable system and workload models*, Adv. in Appl. Probab. **36** (2004), no. 2, 471–498. MR 2058146 (2005b:60194)

[MM13] L. Miclo and P. Monmarché, *É?tude spectrale minutieuse de processus moins indécis que les autres*, to appear in Séminaire de Probabilités. Preprint arXiv:1209.3588 available on arXiv, 2013.

[MT93] S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*, Communications and Control Engineering Series, Springer-Verlag London Ltd., London, 1993. MR MR1287609 (95j:60103)

[ODA88] H. G. Othmer, S. R. Dunbar, and W. Alt, *Models of dispersal in biological systems*, J. Math. Biol. **26** (1988), no. 3, 263–298. MR 949094 (90a:92064)

[Pag98] G. Pagès, *A space quantization method for numerical integration*, J. Comput. Appl. Math. **89** (1998), no. 1, 1–38.

[PR05] B. Perthame and L. Ryzhik, *Exponential decay for the fragmentation or cell-division equation*, J. Differential Equations **210** (2005), no. 1, 155–177. MR 2114128 (2006b:35328)

[PTW10] K. Pakdaman, M. Thieullen, and G. Wainrib, *Fluid limit theorems for stochastic hybrid systems with application to neuron models*, Adv. in Appl. Probab. **42** (2010), no. 3, 761–794.

[RH83] H. Ramlau-Hansen, *Smoothing counting process intensities by means of kernel functions*, Ann. Statist. **11** (1983), no. 2, 453–466. MR 696058 (85g:62150)

[Vil09] C. Villani, *Hypocoercivity*, Mem. Amer. Math. Soc. **202** (2009), no. 950, iv+141. MR 2562709 (2011e:35381)

[Wan10] F.-Y. Wang, *Harnack inequalities on manifolds with boundary and applications*, J. Math. Pures Appl. (9) **94** (2010), no. 3, 304–321. MR 2679029 (2011f:58040)