A unifying formulation of the Fokker-Planck-Kolmogorov equation for general stochastic hybrid systems

Julien Bect

A unifying formulation of the Fokker-Planck-Kolmogorov equation for general stochastic hybrid systems

Julien Bect

Department of Signal Processing and Electronic Systems, Supelec, Gif-sur-Yvette, France. E-mail: julien.bect@supelec.fr

Abstract: A general formulation of the Fokker-Planck-Kolmogorov (FPK) equation for stochastic hybrid systems is presented, within the framework of Generalized Stochastic Hybrid Systems (GSHS). The FPK equation describes the time evolution of the probability law of the hybrid state. Our derivation is based on the concept of mean jump intensity, which is related to both the usual stochastic intensity (in the case of spontaneous jumps) and the notion of probability current (in the case of forced jumps). This work unifies all previously known instances of the FPK equation for stochastic hybrid systems, and provides GSHS practitioners with a tool to derive the correct evolution equation for the probability law of the state in any given example.

Keywords: general stochastic hybrid systems, Markov models, continuous-time Markov processes, jump processes, Fokker-Planck-Kolmogorov equation, generalized Fokker-Planck equation

1. INTRODUCTION

Among all continuous-time stochastic models of (nonlinear) dynamical systems, those with the Markov property are especially appealling because of their numerous nice properties. In particular, they come equipped with a pair of operator semigroups, the so-called backward and forward semigroups, which are the analytical keys to most practical problems involving Markov processes. When the system is determined by a stochastic differential equation, these semigroups are generated by Partial Differential Equations (PDE) — respectively the backward and forward Kolmogorov equations. The forward Kolmogorov PDE, also known as the Fokker-Planck equation, rules the time evolution $t \mapsto \mu_t$, where μ_t is the probability distribution of the state X_t of the system at time t. This paper deals with the generalization of this Fokker-Planck-Kolmogorov (FPK) equation to the framework of General Stochastic Hybrid Systems (GSHS) recently proposed by Bujorianu and Lygeros (2004, 2006).

The GSHS framework encompasses nearly all continuous-time Markov models arising in practical applications, including piecewise deterministic Markov processes (Davis, 1984, 1993) and switching diffusions (Ghosh et al., 1992, 1997). Two kinds of jumps are allowed in a GSHS: spontaneous jumps, defined by a state-dependent stochastic intensity $\lambda(X_t)$, and forced jumps triggered by a so-called guard set G. Generalized FPK equations have been given in the literature, in the case of spontaneous jumps, for several classes of models; see Gardiner (1985), Kontorovich and Lyandres (1999), Krystul et al. (2003) and Hespanha (2005) for instance. The case of forced jumps is harder to analyze, at the FPK level, because no stochastic intensity exists for these jumps. Until recently, the only results available in the literature were dealing with one-dimensional models; see Feller (1952, 1954) and Malhamé and Chong (1985). These results have been extended to a class of multi-dimensional models by Bect et al. (2006).

The main contribution of this paper is general formulation of the FPK equation for GSHS's. It is based on the concept of *mean jump intensity*, which conveniently substitutes for the stochastic intensity when the latter does not exist. This equation unifies all previously known instances of the FPK equation for stochastic hybrid systems, and provides GSHS practitioners with a tool to derive the correct evolution equation for the probability law of the state in any given example. The results presented in this paper are extracted from the PhD thesis of the author (Bect, 2007).

The paper is organized as follows. Section 2 introduces our notations for the GSHS formalism, together with various assumptions that will be needed in the sequel. In Section 3 we define the crucial concept of mean jump intensity, which is used in Section 4 to derive our general formulation of the FPK equation for GSHS's. Section 5 concludes the paper with a series of examples and some general remarks concerning PDEs and integro-differential equations.

2. GENERAL STOCHASTIC HYBRID SYSTEMS

The object of interest in the GSHS formalism is a continuous-time strong Markov process $X = (X_t)_{t \geq 0}$, with values in a metric space E^0. It is defined on a filtered space $(\Omega, \mathcal{A}, \mathcal{F})$, equipped with a system $\{ \boldsymbol{P}_x ;\ x \in E^0 \}$ of probability measures

on $(\Omega, \mathcal{A})$, with the property that $\boldsymbol{P}_x\{X_0 = x\} = 1$ for all $x \in E^0$ (i.e., X starts from x under $\boldsymbol{P}_x$). As usual, $\boldsymbol{E}_x$ denotes the expectation operator corresponding to $\boldsymbol{P}_x$.

It is assumed that, for each $\omega \in \Omega$, the samplepath $t \mapsto X_t(\omega)$ is right-continuous, has left limits $X_t^-(\omega)$ in the completion E of E^0, and has a finite number of jumps, denoted by $N_t(\omega)$, on the interval $[0;t]$ for all $t \geq 0$. The last condition can be seen as a "pathwise non-Zenoness" requirement. We will denote by τ_k the k^{th} jump time, with $\tau_k = +\infty$ if there is less than k jumps.

2.1 The hybrid state space

The (completed) state-space of the model is assumed to have a hybrid structure: $E = \cup_{q\in\mathcal{Q}} \{q\} \times E_q$, where $\mathcal{Q}$ is a finite or countable set, and each E_q is either the closure of some connected open subset $D_q \subset \mathrm{R}^{n_q}$ ($n_q \geq 1$) or any singleton space (in which case we set $n_q = 0$). The state at time t can therefore be written as a pair $X_t = (Q_t, Z_t)$, where $Q_t \in \mathcal{Q}$ and $Z_t \in E_{Q_t}$. We denote by $\mathcal{Q}^{\mathrm{d}} = \{q \in \mathcal{Q} \mid n_q = 0\}$ the set of all "purely discrete" modes, and by $E^{\mathrm{d}} = \cup_{q\in\mathcal{Q}^{\mathrm{d}}} \{q\} \times E_q$ the corresponding subset of E.

The state space E is regarded as the disjoint sum of the sets E_q, $q \in \mathcal{Q}$, and endowed with the disjoint union topology [1]. We denote by $\mathcal{E}$ the Borel σ-algebra, and by $\mathcal{E}_{\mathrm{c}}$ the subsets of all relatively compact $\Gamma \in \mathcal{E}$. Moreover, we define a "volume measure" on E by the relation

$$\mathfrak{m}(\Gamma) = \sum_{q\notin\mathcal{Q}^{\mathrm{d}}} \mathfrak{m}_q(\Gamma \cap E_q) + \sum_{x\in E^{\mathrm{d}}} \delta_x(\Gamma)\,, \quad \Gamma \in \mathcal{E},$$

where $\mathfrak{m}_q$ is the n_q-dimensional Lebesgue measure on E_q and δ_x the Dirac mass at x. (Note that $E_q \subset \mathrm{R}^{n_q}$ has been tacitly identified with $\{q\} \times E_q \subset E$.)

Let ∂E_q be the boundary of E_q in R^{n_q}, with the convention that $\partial E_q = \varnothing$ when $n_q = 0$. We define the boundary ∂E of the state space by the relation $\partial E = \cup_{q\in\mathcal{Q}} \{q\} \times \partial E_q$, and the *guard set* by $G = E \setminus E^0$. It is *not* required that $G = \partial E$.

Notations. Let $\mu : \mathcal{E} \to \mathrm{R}$ be a (signed) measure, $K : E \times \mathcal{E} \mapsto \mathrm{R}$ a kernel and $\varphi : E \to \mathrm{R}$ a measurable function. The following notations will be used throughout the paper, assuming the integrals exist: $(\mu K)(\mathrm{d}y) = \int \mu(\mathrm{d}x)\, K(x, \mathrm{d}y)$, $(K\varphi)(x) = \int K(x, \mathrm{d}y)\, \varphi(y)$ and $\mu\varphi = \int \mu(\mathrm{d}x)\, \varphi(x)$.

2.2 A stochastic differential equation with jumps

A vector field $\mathbf{g}$ on E is regarded as a first order differential operator with respect to the continuous variables: its action on a continuously differentiable function $\varphi \in C^1(E)$ will be denoted by $\mathbf{g}\varphi$, where $(\mathbf{g}\varphi)(q,z) = \sum_{i=1}^{n_q} \mathbf{g}^i(q,z) \frac{\mathrm{d}\varphi}{\mathrm{d}z^i}(q,z)$ on $E \setminus E^{\mathrm{d}}$ and $\mathbf{g}\varphi = 0$ on E^{d}. The number of "components" of $\mathbf{g}$ depends on the mode q: to simplify the notations, we shall agree that the indexes i and j always correspond to summations on the number of continuous variables, and drop the explicit dependence on q. For instance, the definition of $\mathbf{g}\varphi$ can be rewritten as $\mathbf{g}\varphi = \sum_i \mathbf{g}^i \frac{\partial\varphi}{\partial z^i}$.

The process X is assumed to be driven by an Itô stochastic differential equation between its jumps: there exist $r+1$ smooth vector fields $\mathbf{f}_l$ and a r-dimensional Wiener process B such that, in mode $q \in \mathcal{Q} \setminus \mathcal{Q}^{\mathrm{d}}$,

$$\mathrm{d}Z_t \;=\; \mathbf{f}_0(q, Z_t)\,\mathrm{d}t + \sum_{l=1}^{r} \mathbf{f}_l(q, Z_t)\,\mathrm{d}B_t^l\,. \tag{1}$$

In other words, for all $\varphi \in C^2(E)$, X satisfies the following generalized Itô formula

$$\varphi(X_t) - \varphi(X_0) = \int_0^t (L\varphi)(X_s)\,\mathrm{d}s + \sum_{l=1}^{r} \int_0^t (\mathbf{f}_l\varphi)(X_s)\,\mathrm{d}B_s + \sum_{0<\tau_k\leq t} \left(\varphi(X_{\tau_k}) - \varphi(X_{\tau_k}^-)\right),$$

where L is the differential generator associated with (1), i.e. $L = \sum_i \mathbf{f}_0^i \frac{\partial}{\partial z^i} + \sum_{i,j} \left(\sum_{l=1}^r \mathbf{f}_l^i \mathbf{f}_l^j\right) \frac{\partial^2}{\partial z^i \partial z^j}$. We make the following smoothness assumption:

Assumption 1. *The drift $\mathbf{f}_0$ is of class C^1, and the other vector fields $\mathbf{f}_l$, $1 \leq l \leq r$, are of class C^2.*

2.3 Two different kinds of jumps

We assume that there exists a Markov kernel K from E to E^0 and a measurable locally bounded function $\lambda : E^0 \to \mathrm{R}_+$, such that the following *Lévy system identity* holds for all $x \in E^0$, $t \geq 0$, and for all measurable $\varphi : E \times E^0 \to \mathrm{R}_+$:

$$\boldsymbol{E}_x\left\{\sum\nolimits_{0<\tau_k\leq t} \varphi(X_{\tau_k}^-, X_{\tau_k})\right\} = \boldsymbol{E}_x\left\{\int_0^t (K\varphi)(X_s^-)\,\mathrm{d}H_s\right\}$$

where $(K\varphi)(y) = \int_{E^0} K(y, \mathrm{d}y')\,\varphi(y, y')$ and H is the predictable increasing process defined by

$$H_t = \int_0^t \lambda(X_s)\,\mathrm{d}s + \sum_{\tau_k\leq t} 1_{X_{\tau_k}^-\in G}\,. \tag{2}$$

The first part corresponds to *spontaneous* jumps, triggered "randomly in time" with a stochastic intensity $\lambda(X_t)$, while the other part corresponds to *forced* jumps, triggered when X hits the guard set G.

Remark. The terms "spontaneous" and "forced" seem to have been coined by Bujorianu et al. (2003). They are closely related to the probabilistic notions of predictability and total inaccessibility for stopping times (see, e.g., Rogers and Williams, 2000, chapter VI, §§12–18), but be shall not discuss this point further in this paper.

Remark. The pair (K, H) is a *Lévy system* for the process X in the sense of Walsh and Weil (1972, definition 6.1). Most authors require that H be continuous in the definition of a Lévy system, thereby disallowing predictable jumps.

3. MEAN JUMP INTENSITY

From now on, we assume that some initial probability law μ_0 has been chosen, with $\mu_0(G) = 0$ since the process cannot start from G. All expectations will be taken, without further mention, with respect to the probability $\boldsymbol{P}_{\mu_0} = \int \mu_0(\mathrm{d}x)\boldsymbol{P}_x$.

[1] which is (here) locally compact, separable and completely metrizable

3.1 Definition and link with the usual stochastic intensity

It is assumed from now on that $\boldsymbol{E}(N_t) < +\infty$. This is a usual requirement for stochastic hybrid processes [2], which is clearly stronger than piecewise-continuity of the samplepaths. Its being satisfied depends not only on the dynamics of the system but also on the initial probability law μ_0.

In order to introduce the main concept of this section, let us define a (positive, unbounded) measure R on $E \times (0;+\infty)$ by

$$R(A) = \boldsymbol{E}_{\mu_0}\left\{\sum\nolimits_{k\geq 1} 1_A\left(X_{\tau_k}^-, \tau_k\right)\right\}.$$

For any $\Gamma \in \mathcal{E}$, the quantity $R(\Gamma \times (0;t])$ is the expected number of jumps starting from Γ during the time interval $(0;t]$.

Definition 2. Suppose that there exists a mapping $r : t \mapsto r_t$, from $[0;+\infty)$ to the set of all positive bounded measures on E, such that, for all $\Gamma \in \mathcal{E}$,

(1) $t \mapsto r_t(\Gamma)$ is measurable,
(2) for all $t \geq 0$, $R(\Gamma \times (0;t]) = \int_0^t r_s(\Gamma)\,\mathrm{d}s$.

Then r is called the *mean jump intensity* of the process X (started with the initial law μ_0).

Let us split R into the sum of two measures R^0 and R^G, corresponding respectively to the spontaneous and forced jumps of the process. Then, using the Lévy system identity, it is easy to see that a mean jump intensity r^0 always exist for the spontaneous part R^0: it is given by

$$r_t^0(\Gamma) = \boldsymbol{E}\big(\lambda(X_t)\,1_{X_t\in\Gamma}\big) = \int_\Gamma \lambda(x)\,\mu_t(\mathrm{d}x)\,.$$

In other words: for spontaneous jumps, a mean jump intensity always exists, and it is the expectation of the stochastic jump intensity $\lambda(X_t)$ on the event $\{X_t \in \Gamma\}$.

Forced jumps are more problematic. The Lévy system identity is powerless here, since no stochastic intensity exists (because forced jumps are predictable). All hope is not lost, though: a simple example will be presented in the next subsection, proving that a mean jump intensity can exist anyway. This is fortunate, since the existence of a mean jump intensity will be an essential ingredient for our unified formulation of the generalized FPK equation. See subsection 5.2 for further details on that issue.

3.2 Where μ_0 comes into play: an illustrative example

Consider the following hybrid dynamics on $E = [0;1]$: the state X_t moves to the right at constant speed $v > 0$ as long as it is in $E^0 = [0;1)$, and jumps instantaneously to 0 as soon as it hits the guard $G = \{1\}$ (i.e., the reset kernel is such that $K(1,\cdot) = \delta_0$).

If we take $\mu_0 = \delta_0$ for the initial law, then the process jumps from 1 to 0 each time t is a multiple of $1/v$, i.e. $\tau_k = k/v$ and $X_{\tau_k}^- = 1$ almost surely. There is therefore no mean jump intensity in this case, since $R = \sum_{k\geq 1} \delta_{(1,\,k/v)}$.

[2] See, e.g., Davis (1984) or Bujorianu and Lygeros (2004).

Now take μ_0 to be the uniform probability on $[0;1]$ (which is, incidentally, the only stationary probability law of the process). Then

$$\begin{aligned} R\big(\Gamma\times(0;t]\big) &= \delta_1(\Gamma)\int_0^1 \operatorname*{argmax}_{k\geq 1}\left\{\frac{k-x}{v}\leq t\right\}\mathrm{d}x \\ &= \delta_1(\Gamma)\int_0^1 \lceil vt+x\rceil\;\mathrm{d}x \\ &= vt\,\delta_1(\Gamma)\,, \end{aligned}$$

where $\lceil vt+x\rceil$ is the smaller integer greater or equal to $vt+x$. Therefore the mean jump intensity exists in this case, and is equal to $v\,\delta_1$ (it is of course time-independent, since μ_0 is stationary). In particular, the global mean jump intensity is $r_t(E) = v$.

4. MAIN RESULTS

4.1 A weak form of the Fokker-Planck-Kolmogorov equation

Taking expectations in 2.2, the following *generalized Dynkin formula* is obtained: for all compactly supported $\varphi \in C^2(E)$ and all $t \geq 0$,

$$\begin{aligned} \boldsymbol{E}\left\{\varphi(X_t)-\varphi(X_0)\right\} &= \boldsymbol{E}\left\{\int_0^t (L\varphi)(X_s)\,\mathrm{d}s\right\} \\ &\quad + \boldsymbol{E}\left\{\sum_{0<\tau_k\leq t}\varphi(X_{\tau_k})-\varphi(X_{\tau_k}^-)\right\}. \end{aligned} \tag{3}$$

Let us assume the existence of a mean jump intensity r_t at all times. Then (3) can be rewritten as

$$(\mu_t-\mu_0)\,\varphi = \int_0^t \mu_s(L\varphi)\,\mathrm{d}s + \int_0^t r_s(K-I)\varphi\,\mathrm{d}s\,, \tag{4}$$

where μ_t is the law of X_t and I is the "identity kernel" on E, i.e. the kernel defined by $I(y,\mathrm{d}y') = \delta_y(\mathrm{d}y')$. Formally differentiating (4) yields

$$\mu_t' = L^*\mu_t + r_t(K-I)\,, \tag{5}$$

where $t \mapsto \mu_t'$ is the "derivative" of $t \mapsto \mu_t$ (in a sense to be specified later), and L^* the adjoint of L in the sense of distribution theory.

Equation (5) begins like the usual Fokker-Planck equation for diffusion processes ($\mu_t' = L^*\mu_t$) and ends with an additional term that accounts for the jumps of the process. Accordingly,

Definition 3. we will say that $t \mapsto \mu_t$ is a solution in the weak sense of the *generalized FPK equation* for the GSHS if

a) there exists a mean jump intensity $t \mapsto r_t$,
b) there exists a mapping $t \mapsto \mu_t'$, from $[0;+\infty)$ to the space $\mathcal{M}_{\mathrm{c}}(E)$ of all Radon measures on E, such that $t \mapsto \mu_t(\Gamma)$ is absolutely continuous with a.e.-derivative $t \mapsto \mu_t'(\Gamma)$, for all $\Gamma \in \mathcal{E}_{\mathrm{c}}$,
c) $L^*\mu_t$ is a Radon measure for all $t \geq 0$,
d) equation (5) holds as an equality between Radon measures, i.e. $\mu_t'(\Gamma) = (L^*\mu_t)(\Gamma) + r_t(K-I)(\Gamma)$ for all $t \geq 0$ and all $\Gamma \in \mathcal{E}_{\mathrm{c}}$.

Such a weak form of the FPK equation is the price to pay for a unified treatment of both kind of jumps. Conditions 3.a and 3.b can be seen as smoothness requirements with respect to the time variable, and 3.c with respect to the space variables.

4.2 "Physical" interpretation

The usual FPK equation admits a well-known physical interpretation as a conservation equation for the "probability mass" (see e.g. Gardiner, 1985). Indeed, assuming the existence of a smooth pdf $p \in C^{2,1}(E \times \mathrm{R}_+)$, the equation $\mu'_t = L^*\mu_t$ can be rewritten as a conservation equation $\partial p_t/\partial t = \mathrm{div}(\mathbf{j}_t)$, with the *probability current* $\mathbf{j}_t$ defined by

$$\mathbf{j}_t^i = \mathbf{f}_0^i\, p_t - \frac{1}{2}\sum_j \frac{\partial(a^{ij}p_t)}{\partial z^j}\,, \quad a^{ij} = \sum_{l=1}^{r} \mathbf{f}_l^i \mathbf{f}_l^j\,. \qquad (6)$$

The additional "jump term" in the generalized FPK equation, admit a nice physical interpretation as well. To see this, let us rewrite it as the difference of two bounded positive measure: $r_t(K-I) = r_t^{\mathrm{src}} - r_t$, where $r_t^{\mathrm{src}} = r_t K$. Therefore r_t and r_t^{src} behave respectively as a *sink* and a *source* in the generalized FPK equation: for each $\Gamma \in \mathcal{E}$, $r_t(\Gamma)\,\mathrm{d}t$ is the probability mass leaving the set Γ during $\mathrm{d}t$, because of the jumps of the process, while $r_t^{\mathrm{src}}(\Gamma)\,\mathrm{d}t$ is the probability mass entering Γ.

These two measures are in fact connected by the reset kernel $K(x, \mathrm{d}y)$. In particular, the relation $r_t(E) = r_t^{\mathrm{src}}(E)$ holds at all times $t \geq 0$, ensuring that the total probability mass is conserved. Moreover, introducing the measures $W_t(\mathrm{d}x, \mathrm{d}y) = r_t(\mathrm{d}x)K(x, \mathrm{d}y)$, we have $r_t = \int W(\cdot, \mathrm{d}x)$, $r_t^{\mathrm{src}} = \int W(\mathrm{d}x, \cdot)$ and the generalized FPK equation can be rewritten more symmetrically as

$$\mu'_t \;=\; L^*\mu_t + \int \left(W_t(\mathrm{d}x, \cdot) - W_t(\cdot, \mathrm{d}x)\right)\,.$$

It appears clearly, under this form, as a generalization of the *differential Chapman-Kolmogorov formula* of Gardiner (1985, equation 3.4.22) — which only allows spontaneous jumps.

4.3 Main theorem

The main result of this paper show that the various requirements of definition 3 are not independent. We denote by $|\nu|$ the total variation measure of a Radon measure ν, which is finite on $\mathcal{E}_{\mathrm{c}}$. We shall say that a function $t \mapsto \nu_t$ from $[0; \infty)$ to $\mathcal{M}_{\mathrm{c}}(E)$ is right-continuous (resp. locally integrable) is $t \mapsto \nu_t\varphi$ is right-continuous (resp. locally integrable) for all bounded measurable $\varphi : E \to \mathrm{R}$.

Theorem 4. *Consider the following assumptions:*

- *a) there exists a mean jump intensity r (3.a), such that $t \mapsto r_t$ is right-continuous,*
- *b) $t \mapsto \mu_t$ is differentiable in the sense of 3.b, $t \mapsto \mu'_t$ is right-continuous and $t \mapsto |\mu'_t|$ locally integrable,*
- *c) $L^*\mu_t$ is a Radon measure for all $t \geq 0$ (3.c), $t \mapsto L^*\mu_t$ is right-continuous and $t \mapsto |L^*\mu_t|$ is locally integrable.*

If any two of these assumptions hold, then the third holds as well and $t \mapsto \mu_t$ is a solution in the weak sense of the generalized FPK equation.

The proof of this theorem is given in appendix A. We will not try to give general conditions under which assumptions 4.a–4.c are satisfied, since such conditions would inevitably be, in the general setting of this paper, very complicated (involving the initial law μ_0, the vector fields $\mathbf{g}$ of the stochastic differential equation, the geometry of the state space E and the reset kernel K).

4.4 The case when a piecewise smooth pdf exists

Equation (5) is an evolution equation for the measure-valued function $t \mapsto \mu_t$. In most situations of practical interest, the measures μ_t admit a pdf p_t, with respect to the volume measure $\mathfrak{m}$ on E (sometimes with an additional singular measure, like a linear combination of Dirac masses, but this case will not be discussed here). If the function $p : (x, t) \mapsto p_t(x)$ is smooth enough, at least piecewise, then equation (5) simultaneously gives birth to an evolution equation for $t \mapsto p_t$ and to static relations that hold for all $t \geq 0$ (so-called "boundary conditions", although the name is not entirely appropriate here). This can be done quite generally, using some additional measure-theoretic tools for which there is no room in this paper. The reader is referred to Bect (2007, §IV.2.C) for more on this issue.

5. EXAMPLES

5.1 A class of models with spontaneous jumps

Our first series of examples covers a large family of models without forced jumps ($G = \varnothing$). The reset kernel K is assumed to satisfy the following assumption:

Assumption 5. *There exists a kernel K^* on E such that*

$$\mathfrak{m}(\mathrm{d}x)\, K(x, \mathrm{d}y) = \mathfrak{m}(\mathrm{d}y)\, K^*(y, \mathrm{d}x)\,.$$

(We do *not* assume that K^* is a Markov kernel, i.e. that $K^*(y, \cdot)$ is a probability measure for all y.) The following result is an easy consequence of Theorem 4:

Corollary 6. *If there exists a pdf $p \in C^{2,1}(E \times \mathrm{R}_+)$, then the measures r_t and r_t^{src} are absolutely continuous with respect to $\mathfrak{m}$,*

$$\frac{\mathrm{d}r_t}{\mathrm{d}\mathfrak{m}} \;=\; \lambda\, p_t\,, \quad \frac{\mathrm{d}r_t^{\mathrm{src}}}{\mathrm{d}\mathfrak{m}} \;=\; K^*\left(\lambda\, p_t\right)\,,$$

and the following evolution equation holds:

$$\frac{\partial p_t}{\partial t} \;=\; L^* p_t \;+\; K^*\left(\lambda\, p_t\right) \;-\; \lambda\, p_t\,. \qquad (7)$$

Assumption 5 holds for several classes of models known in the literature: pure jump processes with an absolutely continuous reset kernel, the switching diffusions of Ghosh et al. (1992, 1997) and also the SHS of Hespanha (2005).

Example 7. Pure jump processes occur when $L = 0$, i.e. when there is no continuous dynamics. We consider here the case where K is absolutely continuous: $K(x, \mathrm{d}y) = k(x, y)\,\mathfrak{m}(\mathrm{d}y)$. For instance, if the amplitude of the jumps is independent of the pre-jump state and distributed the pdf ρ, then $k(x, y) = \rho(y - x)$. In this case Assumption 5 holds with $K^*(x, \mathrm{d}y) = k(y, x)\,\mathfrak{m}(\mathrm{d}y)$. Introducing the function $\gamma(x, y) = \lambda(x)k(x, y)$, equation 7 turns into the well-known *master equation* (Gardiner, 1985, eq. 3.5.2):

$$\frac{\partial p}{\partial t}(y, t) \;=\; \int \big(\gamma(x, y)p(x, t) - \gamma(y, x)p(y, t)\big)\,\mathfrak{m}(\mathrm{d}x)\,.$$

In particular, when all modes are purely discrete ($n_q = 0$), this is just the usual forward Kolmogorov equation for a continuous-time Markov chain.

Example 8. In the case of switching diffusions, the state space is of the form $E = \mathcal{Q} \times \mathrm{R}^n$ (with $\mathcal{Q}$ a countable set and $n \geq 1$) and the reset kernel of the form

$$K\big((q,z),\cdot\big) = \sum_{q' \neq q} \pi_{qq'}(z)\, \delta_{(q',z)}\,,$$

where $\pi(z) = (\pi_{qq'}(z))$ is a stochastic matrix for all $z \in \mathrm{R}^n$. Assumption 5 is fulfilled with K^* defined by

$$K^*\big((q,z),\cdot\big) = \sum_{q' \neq q} \pi_{q'q}(z)\, \delta_{(q',z)}\,.$$

Equation 6 becomes in this case the familiar generalized FPK equation for switching diffusion processes (see, e.g., Kontorovich and Lyandres, 1999; Krystul et al., 2003): for all $x = (q,z) \in E$ and $t \geq 0$,

$$\frac{\partial p}{\partial t}(x,t) = (L^* p_t)(x) + \sum_{q' \neq q} \lambda_{q'q}(z)\, p_t(q',z) - \lambda(x)\, p_t(x)\,,$$

where $\lambda_{q'q}(z) = \lambda(q',z)\, \pi_{q'q}(z)$.

Example 9. The SHS of Hespanha (2005) are also defined on $E = \mathcal{Q} \times \mathrm{R}^n$, but this time the post-jump state X_{τ_k} is determined by applying a reset map $\Psi : E \to E^0$ to the pre-jump state $X_{\tau_k}^-$, Ψ being chosen randomly in a finite of reset maps Ψ_k. The reset kernel can therefore be written as

$$K(x,\cdot) = \sum_k \pi_k(x)\, \delta_{\Psi_k(x)}\,,$$

with $\pi_k(x)$ the probability of choosing the reset map Ψ_k given that $X_{\tau_k}^- = x$. Provided that the functions Ψ_k are local C^1-diffeomorphisms, the kernel K fulfills Assumption 5 with

$$K^*(x,\cdot) = \sum_k \sum_{y \in \Psi_k^{-1}(\{x\})} \pi_k(y)\, \big|J_k(y)\big|^{-1}\, \delta_y\,,$$

where $J_k(y)$ is the Jacobian determinant of Ψ_k at y. Therefore, introducing a stochastic intensity $\lambda_k = \lambda\, \varrho_k$ for each one of the reset maps, we recover thanks to Corollary 6 the generalized FPK equation given by Hespanha (2005, p. 1364):

$$\begin{aligned}\frac{\partial p}{\partial t}(x,t) \;&=\; (L^* p_t)(x) \\ &\quad + \sum_k \sum_{y \in \Psi_k^{-1}(\{x\})} \left(\frac{\lambda_k\, p_t}{|J_k|}(y) \;-\; (\lambda_k\, p_t)(x) \right)\,.\end{aligned}$$

5.2 A class of models with forced jumps

The measure-valued formulation of the generalized FPK equation equation (5) paves the way for an easier proof of some recent results (Bect et al., 2006), concerning GSHS with forced jumps and deterministic resets. A typical example of this class of process is the thermostat model of Malhamé and Chong (1985). Since a complete statement and proof of these results would be too long for this paper, we shall only provide an illustrative example. The interested reader is referred to the PhD thesis of the author (Bect, 2007, IV.2.C and IV.3.C). A thorough treatment will appear in a forthcoming publication.

Example 10. Let us consider a GSHS without spontaneous jumps ($\lambda = 0$), whose hybrid state space is defined by $\mathcal{Q} = \{0,1\}$, $E_0 = [z_{\min}; +\infty) \times \mathrm{R}^{n-1}$, and $E_0 = (-\infty; z_{\max}] \times \mathrm{R}^{n-1}$ (where $z_{\min} < z_{\max}$). Assume that the guard G is the whole boundary ∂E, and that the reset map is defined by $\Psi(q,z) = (1-q,z)$. In other words, the discrete component Q_t switches from 0 to 1 when Z_t^1 reaches the lower threshold $z_{\min}$, and switches back to 0 when Z_t^1 reaches the upper threshold $z_{\max}$.

For such a hybrid structure, it is easily shown using Theorem 4 that no $C^{2,1}$ solution can exist. Consider the set $G' = \Psi(G)$, which is the disjoint unions of two "hyperplanes" in E^0. A careful examination of (5) suggests to look for solution that are of class $C^{2,1}$ on $E^0 \setminus G'$, possibly with a discontinuity on G'. If the process effectively has a pdf p satisfying these assumptions, then it can be proved using Theorem 4 that:

(1) The usual Fokker-Planck equation, $\partial p_t/\partial t = L^* p_t$, holds on the four components of $E^0 \setminus G'$,
(2) The jumps are accounted for by the static relation $j_t^{\mathrm{out}} = j_t^{\mathrm{in}} \circ \psi$ on G, at all times $t \geq 0$, where j_t^{out} and j_t^{in} are the outgoing and ingoing probability current, respectively defined on G and G' (see (6) for the defintion of the probability current).
(3) The mean jump intensity r_t is supported by G and given by the outgoing flux of the probabily current $\mathbf{j}_t$, i.e. $r_t(\Gamma) = \int_{\Gamma \cap G} j_t^{\mathrm{out}} \mathrm{d}\mathfrak{s}$, where $\mathfrak{s}$ is the surface measure.
(4) Finally, for each $x \in G$ such that at least one of the "noise driven" vector fields $\mathbf{g}_l$ ($1 \leq l \leq r$) is transverse to G, the pdf has to satisfy the so-called *absorbing boundary condition* $p_t(x) = 0$. For similar reasons, p_t has to be continuous at each $x \in \Gamma$ such that at least one of the "noise driven" vector fields is transverse to G'.

5.3 A remark concerning PDEs

Notations can be deceiving, sometimes. The compact formulation of (5) and (7), which makes them look very much like the usual Fokker-Planck equation, should not fool the reader into thinking that these equations are simple PDEs. Indeed, even when a (piecewise) smooth pdf exists, the generalized FPK equation is in general a system of integro-differential equations, with boundary conditions that can also involve integrals. The integrals are hidden in the kernel notation: $(r_t K)(\Gamma) = \int r_t(\mathrm{d}x) K(x,\Gamma)$. Fortunately, they disappear in many interesting examples where the reset kernel is simple enough (see examples 8–10). This is an important observation for practical applications, since the numerical solution of a PDE is much easier than that of a general integro-differential equation.

Appendix A. PROOF OF THEOREM 4

Let $C_{\mathrm{c}}^2(E)$ denote the set of all compactly supported $\varphi \in C^2(E)$. The following lemma is an easy consequence of the smoothness of the vector fields:

Lemma 11. *For all $\varphi \in C^2(E)$, $t \mapsto \int_0^t (L^* \mu_s)(\varphi)\, \mathrm{d}s$ is differentiable on the right, with the right continuous derivative $t \mapsto (L^* \mu_t)(\varphi)$.*

In the sequel, "right continuous" is abbreviated as "rc".

$\diamond$ Assume that both 4.a and 4.b hold. Then each term of (4) has a t-derivative on the right. Differentiating both sides proves that (5) holds for all $t \geq 0$, hence that $L^*\mu_t$ is a Radon measure and that $t \mapsto L^*\mu_t$ is rc. Moreover, integrating the inequality $|L^*\mu_t| \leq |\mu'_t| + 2r_t$ yields that, for all $\Gamma \in \mathcal{E}_{\rm c}$,

$$\int_0^t |L^*\mu_s|\,(\Gamma)\,\mathrm{d}s \leq \int_0^t |\mu'_s|\,(\Gamma)\,\mathrm{d}s + 2\,\boldsymbol{E}\big\{N_t\big\} \leq +\infty\,.$$

Therefore $t \mapsto |L^*\mu_s|$ is locally integrable, which proves 4.c.

$\diamond$ Assume now that 4.a and 4.c hold, and set $\mu'_t = L^*\mu_t + r_t(K-I)$, for all $t \geq 0$. Clearly, μ'_t is a Radon measure, $t \mapsto \mu'_t$ is rc and

$$\int_0^t \mu'_t\varphi = (\mu_t - \mu_0)\varphi\,, \quad \forall t \geq 0\,, \quad \forall \varphi \in C_{\rm c}^2(E)\,. \quad \text{(A.1)}$$

Moreover, for all $\Gamma \in \mathcal{E}_{\rm c}$,

$$\int_0^t |\mu'_s|\,(\Gamma)\,\mathrm{d}s \leq \int_0^t |L^*\mu_s|\,(\Gamma)\,\mathrm{d}s + 2\,\boldsymbol{E}\big\{N_t\big\} \leq +\infty\,,$$

which shows that $t \mapsto |\mu'_s|$ is locally integrable. Therefore, using standard approximation techniques and a monotone class argument, it can be proved that (A.1) still holds for $\varphi = 1_\Gamma$, $\Gamma \in \mathcal{E}_{\rm c}$, i.e. that $t \mapsto \mu'_t$ is the "derivative" of $t \mapsto \mu_t$ in the sense of definition 3.b.

$\diamond$ Finally, assume that 4.b and 4.c hold. Then, for all $\varphi \in C_{\rm c}^2(E)$, equation (4) can be rewritten as

$$\iint\limits_{G\times]0;t]} \varphi(x)\,\big(R^G(\mathrm{d}x,\mathrm{d}s) - (L^*\mu_s)(\mathrm{d}x)\mathrm{d}s\big) = \iint\limits_{E^0\times]0;t]} \varphi(x)\,\big((R^G K)(\mathrm{d}x,\mathrm{d}s) - \xi_s(\mathrm{d}x)\,\mathrm{d}s\big)\,, \quad \text{(A.2)}$$

where $\xi_s = \mu'_s - \big(L^*\mu_s\big)(E^0 \cap \,\cdot\,) - r^0(K-I)$. The measures R^G and r^0 have been defined in subsection 3.1. Clearly, $\xi_t \in \mathcal{M}_{\rm c}(E)$ and $t \mapsto \xi_t$ is locally integrable. Using once more standard approximation techniques, one can prove that (A.2) still holds when $\varphi = 1_\Gamma$, with Γ a compact subset of G. In this case the right-hand side vanishes, yielding

$$R^G(\Gamma\times]0;t]) \;=\; \int_0^t (L^*\mu_s)(\Gamma)\,\mathrm{d}s\,.$$

Moreover, since $t \mapsto R^G(\Gamma\times]0;t])$ is increasing and $t \mapsto (L^*\mu_t)(\Gamma)$ is rc, we have $(L^*\mu_t)(\Gamma) \geq 0$ for all $t \geq 0$. This allows to extend (A.2) to all $\Gamma \in \mathcal{E}_{\rm c}$, using a monotone class argument, thus proving the existence of a mean jump intensity $r_t^G = (L^*\mu_s)(G \cap \,\cdot\,)$ for the forced jumps.

REFERENCES

J. Bect. *Processus de Markov diffusif par morceaux: outils analytiques et numériques*. PhD thesis, Université Paris Sud XI, 2007. In french.

J. Bect, H. Baili, and G. Fleury. Generalized Fokker-Planck equation for piecewise-diffusion processes with boundary hitting resets. In *17th Int. Symp. on Mathematical Theory of Networks and Systems (MTNS)*, 2006.

M. L. Bujorianu and J. Lygeros. General stochastic hybrid systems: modelling and optimal control. In *Proc. of the 43rd IEEE Conf. on Decision and Control*, vol. 2, pp. 1872–1877, 2004.

M. L. Bujorianu and J. Lygeros. Toward a general theory of stochastic hybrid systems. In *Stochastic Hybrid Systems: Theory and Safety Critical Applications*, vol. 337 of *LNCIS*, pp. 3–30. Springer Verlag, 2006.

M. L. Bujorianu, J. Lygeros, W. Glover, and G. Pola. A stochastic hybrid system modeling framework. Technical Report WP1, Deliv. 1.2, HYBRIDGE (IST-2001-32460), 2003.

M. H. A. Davis. Piecewise deterministic Markov processes: a general class of nondiffusion stochastic models. *J. R. Stat. Soc. Ser. B*, 46:353–388, 1984.

M. H. A. Davis. *Markov Models and Optimization*. Chapman & Hall, London, 1993.

W. Feller. The parabolic differential equations and the associated semi-groups of transformations. *Ann. of Math.*, 55(3): 468–519, 1952.

W. Feller. Diffusion processes in one dimension. *Trans. Amer. Math. Soc.*, 77:1–31, 1954.

C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer-Verlag, 1985.

M. K. Ghosh, A. Arapostathis, and S. I. Marcus. Optimal control of switching diffusions with application to flexible manufacturing systems. *SIAM J. Control Optim.*, 30(6):1–23, 1992.

M. K. Ghosh, A. Arapostathis, and S. I. Marcus. Ergodic control of switching diffusions. *SIAM J. Control Optim.*, 35 (6):1952–1988, 1997.

J. P. Hespanha. A model for stochastic hybrid systems with application to communication networks. *Non Linear Analysis*, 62:1353–1383, 2005.

V. Kontorovich and V. Lyandres. Dynamic systems with random structure: an approach to the generation of nonstationary stochastic processes. *J. Franklin Institute*, 336:939–954, 1999.

J. Krystul, A. Bagchi, and H. A. P. Blom. Risk decomposition and assessment methods. Technical Report D8.1, HYBRIDGE (IST-2001-32460), 2003.

R. Malhamé and C. Chong. Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system. *IEEE Trans. Automat. Control*, 30(9):854–860, 1985.

L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales. Volume 2: Itô Calculus. 2nd edition*. Cambridge University Press, 2000.

J. B. Walsh and M. Weil. Représentation de temps terminaux et applications aux fonctionnelles additives et aux systèmes de Lévy. *Annales scientifiques de l'E.N.S. 4e série*, 5(1):121–155, 1972.