# A Class of Stochastic Hybrid Systems with State-Dependent Switching Noise

Leth, John-Josef; Rasmussen, Jakob Gulddahl; Schiøler, Henrik; Wisniewski, Rafal

*Published in:*
Decision and Control (CDC), 2012 IEEE 51st Annual Conference on

*DOI (link to publication from Publisher):*
10.1109/CDC.2012.6427010

*Publication date:*
2012

*Document Version*
Accepted author manuscript, peer reviewed version

# A Class of Stochastic Hybrid Systems with State-Dependent Switching Noise

John Leth, Jakob G. Rasmussen, Henrik Schioler, and Rafael Wisniewski.

***Abstract*— In this paper, we develop theoretical results based on a proposed method for modeling switching noise for a class of hybrid systems with piecewise linear partitioned state space, and state-depending switching. We devise a stochastic model of such systems, whose global dynamics is governed by a continuous-time stochastic process. The main result of this paper is that we may identify the realizations of the global dynamics with the solutions of a differential inclusion. Hence, an analysis of switched systems with switching noise can be carried out either based on a non-deterministic method via the differential inclusion, or on a stochastic method via the stochastic process. Furthermore, we describe how to construct intensity plots, which provide a quick overview of the behavior of the system. An example is included to illustrate this.**

## I. Introduction

Switching between dynamical systems is often connected with uncertainties generated by disturbances, malfunctions etc. To model such a scenario, one may use the concept of a stochastic hybrid system, see [1], [2], [3] and references therein for examples of how various uncertainties enters a hybrid system.

The systems studied in this paper belongs to the class of stochastic hybrid systems SHS, [3], [1], [4], [5], [2], [6]. Indeed, our setup is a special case of [1], [6], [4], but where these strive for a general study of SHS, we specialize to obtain results designed for a particular subclass. The trade-off is of course that these results seem not to be valid for the more general class studied in [1], [6], [4].

In this paper, we develop theoretical results based on a proposed method for modeling switching noise for a class of hybrid systems with piecewise linear partitioned state space and state-depending switching, as described in [7]. In the sequel, elements of this class will be referred to as switched systems (we remark that the terminology for such systems in the literature is not consistent, e.g., in [8, Chapter 3.3] it is called a "state-dependent switched system", in [9] it is called a "hybrid system", and in [10] it is called a "piecewise affine system").

We will partition a subset (state space) of the $n$-dimensional euclidean space into convex subsets (polyhedral sets), here called cells, with disjoint interior such that each cell has non-empty interior and the boundary of each region is the union of convex subsets (facets). For a family of vector fields defined on the partition (one for each cell), we construct a differential inclusion whose corresponding set-valued map gives a one point set at each point in the interior of each cell and a finite set otherwise. This differential inclusion is used to describe the (global) dynamics of the switched systems.

In a switched system, a shift from one local system to another, may be thought of as deterministic in the sense that it occurs with probability one when a trajectory "hits" a face. Looking towards (engineering) applications and simulations, this constitutes a problem since faces have measure zero and hence they will not be "hit" by any simulated trajectory, e.g., any step size or sampling rate will result in the points of the discretized trajectory being on one side of a face at one time instant and on the other side at the next time instant. Moreover, in many situations, it is desirable to describe the future behavior of the system. Since such predictions usually are based on observations, one is forced to introduce some uncertainty into the model. A way to model the uncertainty related to a shift is to "thicken the corresponding face", i.e., replace each face with some (small) open neighborhood of it.

In this paper, we incorporate the "thickening of faces" into our model of a switched system. More precisely, we construct an open neighborhood around the faces on which there is defined a probability measure. Subsequently, we use this measure to describe the probability of a shift, in such a way that the longer a trajectory stays within the neighborhood, the higher the probability of a shift becomes. Based on this construction, we devise a stochastic model of a switched system with switching noise, whose global dynamics is governed by a continuous-time stochastic process with values in the state space. The main result of this paper is that, on any finite time interval, we may, except for a set of measure zero, identify the realizations of the global dynamics with the solutions of a differential inclusion. Hence, for switched systems with switching noise which can be modeled as described in this paper, an analysis of such systems can be performed in two ways; a stochastic analysis, via the stochastic process which is useful when simulations of such systems are required, and a non-deterministic analysis via the differential inclusion, if a more detailed analysis is required. Furthermore, we describe how one applies the stochastic model to construct intensity plots, which provide a quick overview of the behavior of the system. We end with an example illustrating the use of intensity plots.

This work was supported by the Danish Council for Technology and Innovation.

J. Leth, H. Schioler and Wisniewski are with Department of Electronic Systems, J. G. Rasmussen is with Department of Mathematical Sciences, all at Aalborg University, Fredrik Bajers Vej 7C, 9220 Aalborg, Denmark jjl@es.aau.dk jgr@math.aau.dk hs@es.aau.dk raf@es.aau.dk

## II. Preliminaries

For completeness, we recall various definitions and results from the theory of switched systems with state-dependent switching.

### A. Switched Systems

Before introducing the concept of a switched system, we recall the definitions of a (polyhedral) complex and a (piecewise linear) partition.

Let $\Lambda$ be some index set, and $K = \{P_i\}_{i\in\Lambda}$ be a family of polyhedral sets in $E = \mathbb{R}^n$. We let $|K| = \cup_{i\in\Lambda} P_i$ with the subspace topology inherited from $E$, and call $K$ a (polyhedral) complex if (1) each face of any $P \in K$ is in $K$, (2) $P \cap P'$ is a face of $P$ and $P'$, for any $P, P' \in K$, and (3) each point of $|K|$ has a neighborhood intersecting only finitely many elements of $K$. Condition (3) is only necessary if $\Lambda$ is infinite. For a complex $K$ with index set $\Lambda$, we let $\Lambda^j = \{i \in \Lambda | \dim(P_i) = j\}$ and $K^j = \{P_i \in K \mid i \in \Lambda^j\}$. Elements $P_i \in K$ will be called cells if $i \in \Lambda^n$, and facets if $i \in \Lambda^{n-1}$.

Let $E'$ denote either $E$ or a bounded polyhedral set (i.e., a polytope) in $E$ of dimension $n$. By a (piecewise linear) partition of $E'$, we mean a complex $K$ such that $|K| = E'$. An $n$-dimensional switched system $\mathcal{S}$ with index set $\Lambda$ is then a triple $\mathcal{S} = (E', K, G)$ where the state space $E'$ is either $E = \mathbb{R}^n$ or a bounded polyhedral set in $E$ of dimension $n$, where the complex $K = \{P_i\}_{i\in\Lambda}$ is a (piecewise linear) partition of $E'$, and where $G = \{f_i\}_{i\in\Lambda^n}$ is a family of smooth functions, $f_i : \mathbb{R}^n \to E$, describing the local dynamics of $\mathcal{S}$. The global dynamics of $\mathcal{S}$ is governed by the following differential inclusion

$$x'(t) \in f(x(t)), \tag{1}$$

where the set valued map $f$ is defined by

$$\begin{aligned} f :& E' \to 2^E; \\ & x \mapsto \{v \in E | v = f_i(x) \text{ for all } i \in \Lambda^n \text{ such that } x \in P_i\}, \end{aligned}$$

with $2^E$ the power set of $E$.

We say that $t \mapsto x(t)$ follows system $i$, and refer to this as a local solution if $x'(t) = f_i(x(t))$. We remark that for technical reasons a local solution is not confined to a particular polyhedral set, e.g., a local solution of $x' = f_i(x)$ is allowed to extend beyond the polyhedral set $P_i$.

We recall various solution concepts related to $\mathcal{S}$. For $0 < T \leq \infty$ let $J_T = [0, T)$. By a (Caratheodory) solution at $x_0 \in E'$ to $\mathcal{S}$, we understand an absolutely continuous (AC) function $J_T \to E'$; $t \mapsto x(t)$, which solves the Cauchy problem

$$x'(t) \in f(x(t)) \text{ a.e.}, \quad x(0) = x_0. \tag{2}$$

A relaxed (or Filippov) solution at $x_0 \in E'$ to $\mathcal{S}$ is by definition a solution at $x_0$ to the differential inclusion

$$x'(t) \in f^c(x(t)) \text{ a.e.}, \quad f^c(x) = \text{co}(f(x)), \tag{3}$$

with $\text{co}(f(x))$ the convex hull of $f(x)$. A classical solution at $x_0 \in E'$ to $\mathcal{S}$ is a continuously differentiable function $J_T \to E'$; $t \mapsto x(t)$, which solves the Cauchy problem; $x'(t) \in f(x(t)),\ x(0) = x_0$.

In the sequel, we usually drop the index 0 when we discuss a family of initial conditions, as in Proposition 1 below.

Now, let $T_{E'}(x)$ denote the contingent cone to $E'$ at $x$, i.e., the closure of the convex cone of $E' - \{x\}$. We end this section with the following existence results which easily can be derived from standard results from differential inclusions, e.g from [11].

*Proposition 1:* At any interior point $x$ of $E'$ there exists a relaxed solution at $x$ to $\mathcal{S}$. Moreover, the solution is a classical solution if $x$ is interior to a cell.

*Proposition 2:* For each unbounded $P_i$, with $i \in \Lambda^n$, assume that $f_i(P_i)$ is bounded. Then at any $x \in E'$ there exists a relaxed solution to $\mathcal{S}$ defined on $[0, \infty)$

1) if $\Lambda^n$ is finite, in the case $E' = E$.
2) iff $f^c(x) \cap T_{E'}(x) \neq \emptyset$ for all $x \in E'$, in the case $E' \neq E$.

## III. Switched Systems With Switching Noise

In this section, we devise a stochastic model $\mathcal{S}_{\epsilon h}$ of a switched system $\mathcal{S}$ with switching noise, which can be interpreted as a (stochastic) approximation of $\mathcal{S}$. The switching noise is assumed to have effect only in an $\epsilon$-neighborhood around the switching surfaces (the facets) and the uncertainty related to the occurrence of a shift within this neighborhood is assumed to be governed by an intensity function $h$ as described below.

As we shall see in Section IV, the stochastic model can be used to derive a curve intensity measure giving a tool for analyzing the mean behavior of solutions to switched systems with switching noise.

### A. Switching noise

In order to introduce switching noise, we will start by recalling the intensity function of a random variable, [12]. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $U : \Omega \to \mathbb{R}$ a random variable with differentiable distribution function $F_U : \mathbb{R} \to \mathbb{R}$, and density function $p : \mathbb{R} \to [0, \infty)$.

Assume that the stochastic variable $U$ is distributed on the interval $[-\epsilon, \epsilon]$. Then the distribution of $U$ can also be described by the survivor function $u \mapsto S(u) = 1 - \int_{-\epsilon}^{u} p(v)dv$, i.e., the probability of $U > u$, $\mathbf{P}(U > u)$, or by its intensity (or hazard) function $u \mapsto h(u)$ with

$$h(u) = \frac{p(u)}{S(u)} = \frac{1}{\mathbf{P}(U > u)} \lim_{\Delta \to 0+} \frac{\mathbf{P}(U < u + \Delta) - \mathbf{P}(U < u)}{\Delta}, \tag{4}$$

i.e., $\Delta h(u)$ is, for small $\Delta > 0$, approximately the probability of $U$ being in small interval $(u, u + \Delta)$, conditional on $U$ being bigger than $u$, $\mathbf{P}(U \in (u, u + \Delta) \mid U > u)$. The intensity function turns out to be a convenient starting point for defining the stochastic shifts between the different dynamical systems.

It should be remarked that under mild conditions on $h$, there is a one to one correspondence between $h$ and $p$. More precisely, if $h : [-\epsilon, \epsilon] \to \mathbb{R}$ is any map such that

$$u \mapsto 1 - \exp(-\int_{-\epsilon}^{u} h(v)dv), \quad u \in [-\epsilon, \epsilon],$$

defines a differentiable distribution function, then $h$ is an intensity function with the corresponding density function $p$ given by

$$p(u) = h(u)\exp(-\int_{-\epsilon}^{u} h(v)dv), \quad u \in [-\epsilon, \epsilon]. \tag{5}$$

To obtain an intensity function suitable for our purpose, we need the following, see also Figure 2. Let $\mathcal{S} = (E', K, G)$ denote an $n$-dimensional switched system with index set $\Lambda$ and for each pair of neighboring polyhedral sets, say $P_i$ and $P_j$, define the map $x_{ij} = x_{ij}(\cdot)$ by

$$E' \to \mathbb{R}^n;\ x \mapsto x_{ij} = x - \pi_{\mathbb{F}_{ij}}(x),$$

where $\mathbb{F}_{ij}$ is the affine hull of the facet $F_{ij} = P_i \cap P_j$ between $P_i$ and $P_j$, and $\pi_{\mathbb{F}_{ij}}$ is the orthogonal projector onto $\mathbb{F}_{ij}$, [11, p. 24]. Note that $x_{ij}$ is a vector pointing into $P_i$ (resp. $P_j$) when $x \in P_i$ (resp. $x \in P_j$).

Let $n_{ij}$ (resp. $n_{ji}$) denote a normal vector for the facet $F_{ij}$ which points into $P_j$ (resp. $P_i$), and define the map

$$E' \to \mathbb{R};\ x \mapsto u_{ij}(x) = |x_{ij}|\text{sign}(n_{ij}|x_{ij}), \quad x_{ij} = x_{ij}(x).$$

To define the intensity function $h_{ij}$ for a shift from system $i$ to system $j$, we need the following construction. For $\epsilon > 0$ and $P \in K$, we let $P^\epsilon$ denote the $\epsilon$-expansion of $P$ defined as $P^\epsilon = \cap H_i^{\epsilon,H}$ with $\{H_i\}$ the (finite) family of half spaces defining $P$ (i.e., $P = \cap H_i$) and $H_i^{\epsilon,H}$ the $\epsilon$-neighborhood of $H_i$ with respect to the Hausdorff metric. Note that $P^\epsilon \neq P^{\epsilon,H}$, e.g., if $P$ is a square centered at $c$ with side length $l$ then $P^\epsilon$ is a square centered at $c$ with side length $2\epsilon + l$ and with the same orientation as $P$; whereas, $P^{\epsilon,H}$ is $P^\epsilon$ with rounded corners, see Figure 1.

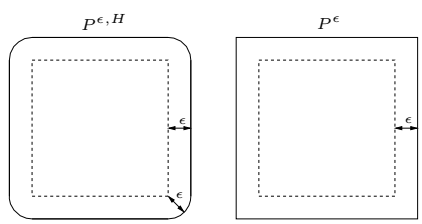



Fig. 1. $P^{\epsilon,H}$ and $P^\epsilon$ with $P$ indicated by the dotted square.

Let $x(\cdot)$ follow system $i$ and $F_{ij} \neq \emptyset$, then the intensity function $h_{ij}$ for a shift from system $i$ to system $j$ at $x(t) \in F_{ij}^\epsilon$ is defined by

$$h_{ij}(x(t)) = h(u_{ij}(x(t))), \tag{6}$$

where $h$ is the intensity function defined above. Hence, $h_{ij}(x(t))$ is the intensity of the signed orthogonal distance from $x(t)$ to $\mathbb{F}_{ij}$, the affine hull of the facet $F_{ij}$ between $P_i$ and $P_j$, i.e., we have used the distribution on $[-\epsilon, \epsilon]$ to induce distributions, one for each solution, on trajectories. We note that for a given solution $x(\cdot)$, the density function connected to $h$ induces a density function connected to $h_{ij}(x(t))$ by (5).

The above construction deals with shifts when a trajectory approaches one facet, but it does not show how to handle

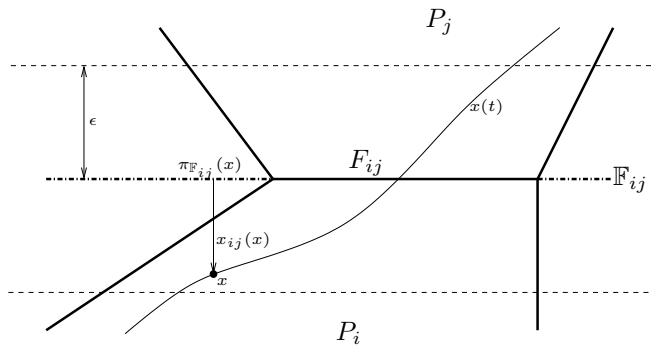



Fig. 2. An $\epsilon$-neighborhood around the affine hull $\mathbb{F}_{ij}$ of the facet $F_{ij}$ between the polyhedral sets $P_i$ and $P_j$ (in the figure four polyhedral sets are shown). Also indicated is the vector $x_{ij}(x) = x - \pi_{\mathbb{F}_{ij}}(x)$, where $\pi_{\mathbb{F}_{ij}}(x)$ is the is the orthogonal projector onto $\mathbb{F}_{ij}$ of the point $x$ on the solution curve $x(\cdot)$.

multiple facets simultaneously. To handle this situation, we assume in sequel that if $x(t) \in \bigcap_{i\in\Lambda'} P_i^\epsilon$ with $\Lambda' \subset \Lambda^n$, then shifts to each of the systems $f_i,\ i \in \Lambda'$ will happen independently.

*Proposition 3:* Assume that $t \mapsto x(t)$ follows system $i$, and that $x(t) \in \bigcap_{i=1}^k P_{j_i}^\epsilon$ with $j_i \in \Lambda^n$. Then the intensity function $h_{i\bullet}$ for a shift from system $i$ to system $l \in \{j_1, \ldots, j_k\}$ at $x(t)$ is given by

$$h_{i\bullet}(x(t)) = \sum_{j=j_1}^{j_k} h_{ij}(x(t)),$$

and if a shift occurs at $x(t)$, we shift to system $l \in \{j_1, \ldots, j_k\}$ with probability

$$h_{il}(x(t))/h_{i\bullet}(x(t)).$$

Equivalently, we can let shifts happen according to all of the intensities $h_{ij_1}(x(t)), \ldots, h_{ij_k}(x(t))$ independently of each other, and disregard all shifts except the first one with respect to time $t$.

*Proof:* Let $x = x(\cdot)$ and $F_x^{ij}$ denote the distribution function corresponding to $h_{ij}$, i.e., $F_x^{ij}(t)$ is the probability of a shift from system $i$ to system $j$ before time $t$. Hence $1 - F_x^{ij}(t)$ is the probability of no shift from system $i$ to system $j$ before time $t$.

Now by assumption shifts to each of the systems will happen independently; hence, we conclude that the product $\Pi_j(1 - F_x^{ij}(t))$ is the probability of no shift from system $i$ before $t$.

Let $F_x^{i\bullet}(\cdot)$ denote the distribution function defined by letting $F_x^{i\bullet}(t)$ be the probability of a shift from system $i$ before $t$. Hence the corresponding intensity function is $h_{i\bullet}$, and we then have

$$\begin{aligned} F_x^{i\bullet}(t) &= 1 - \Pi_j(1 - F_x^{ij}(t)) \\ &= 1 - \Pi_j \exp(-\int_{t'}^{t} h_{ij}(x(t))dt) \\ &= 1 - \exp(-\int_{t'}^{t} \sum_j h_{ij}(x(t))dt), \end{aligned}$$

where $t'$ denotes the time for the last shift or 0 if no shifts have occurred before $t$. Thus $h_{i\bullet}(x(t)) = \sum_{j=j_1}^{j_k} h_{ij}(x(t))$, which proves the first part of the proposition.

Now let $T$ denote the continuous random variable giving the time of the shift from system $i$, and $I$ denote the discrete random variable giving the number of the system which is shifted to. With this notation we use (4) to conclude

$$\begin{aligned}\frac{h_{il}(x(t))}{h_{i\bullet}(x(t))} &= \lim_{\Delta\to 0^+} \frac{\mathbf{P}(T \in (t, t+\Delta), I = l)}{\mathbf{P}(T \in (t, t+\Delta))} \\ &= \lim_{\Delta\to 0^+} \mathbf{P}(I = l \mid T \in (t, t+\Delta)),\end{aligned}$$

hence completing the proof of the last assertion. ■

*Example 1:* Let

$$h(u) = \frac{\mathbb{I}[u \in [-\epsilon, \epsilon]]}{\epsilon - u}, \tag{7}$$

where $\mathbb{I}$ denotes the indicator function. Then by (5) the density function corresponding to $h$ is

$$p(u) = \frac{\mathbb{I}[u \in [-\epsilon, \epsilon]]}{2\epsilon},$$

i.e., $p$ is the density function for the uniform distribution on $[-\epsilon, \epsilon]$. Moreover,

$$h_{ij}(x) = \frac{\mathbb{I}[u_{ij}(x) \in [-\epsilon, \epsilon]]}{\epsilon - u_{ij}(x)},$$

and by (7) any trajectory whose intersection with $P_i^\epsilon \cap P_j^\epsilon$ is contained in a normal subspace to $F_{ij}$ will shift according to the uniform distribution on this part of the trajectory.

### *B. A stochastic model of a switched systems with switching noise*

Having formalized the notion of a shift, we can proceed to the model of a switched system with switching noise.

An $n$-dimensional stochastic switched system with index set $\Lambda$ is a quintuple $\mathcal{S}_{\epsilon h} = (E', K, G, \epsilon, h)$, where $\mathcal{S} = (E', K, G)$ is an $n$-dimensional switched system with index set $\Lambda$, $\epsilon > 0$ is a parameter determining the size of the switching neighborhood, and $h : [-\epsilon, \epsilon] \to (0, \infty)$ is a strictly positive intensity function. Note that we in the definition of a stochastic switched system assume that the intensity function $h$ is strictly positive, a fact that is used to prove Proposition 4.

To define the notion of a solution of a stochastic switched system, we will henceforth assume that any local solution can be extended to $[0, \infty)$, that is, local solutions do not blow up. Moreover, we refer to the part of a local solution, which follow system $i$, say, lying in $P_i^\epsilon$ as an $\epsilon$-local solution.

Now, let $\Omega_{x_0 i_0} = \Omega_{x_0 i_0}(\mathcal{S}_{\epsilon h}) \subset C([0, \infty), E')$ denote the subset of continuous curves from $[0, \infty)$ to $E'$ starting at $x_0$ in system $i_0$ and which are piecewise $\epsilon$-local solutions. Hence, if $x(\cdot) \in \Omega_{x_0 i_0}$, then it starts as a local solution at $x_0$, for system $i_0$ and proceeds according to system $i_0$ and either (1) stays in that system for all time, or (2) shifts to another local solution, dictated by the intensity function $h_{i\bullet}$, in which case it proceeds according to the new system until either case (1) or case (2) occurs.

By a solution to $\mathcal{S}_{\epsilon h}$ at $x_0 \in E'$ (starting in system $i_0 \in \Lambda$), we understand an $E'$-valued continuous-time stochastic process $X = X^{x_0 i_0} : [0, \infty) \times \Omega_{x_0 i_0} \to E'$ given by

$$X(t, x(\cdot)) = X_t(x(\cdot)) = x(t), \quad X_0(x(\cdot)) = x_0 \ \forall x(\cdot)$$

started at $x_0$ which evolves according to the local dynamics $f_i \in G$ and shifts between these systems according to the intensity functions $h_{ij}$ as described in Section III-A.

To a stochastic switched system $\mathcal{S}_{\epsilon h}$ there corresponds a probability space $(\Omega_h, \mathcal{F}_h, \mathbf{P}_h)$ connected to the intensity function $h$ (see (5) and the paragraph above this) and a family of induced probability spaces $\{(\Omega_{xi}(\mathcal{S}_{\epsilon h}), \mathcal{F}_{xi}, \mathbf{P}_{xi})\}_{(x,i)\in\mathcal{A}}$, $\mathcal{A} \subset E' \times \Lambda$ connected to the family of solutions $\{X = X^{xi}\}_{(x,i)\in\mathcal{A}}$ with $\mathbf{P}_{xi}(X_t^{xi} \leq y)$ being the probability of a realization being in the set $\{z \in E' \mid z \leq y\}$ at time $t$, where $z \leq y$ is to be understood coordinate wise.

Moreover, to $X = X^{x_0 i_0}$ there corresponds a discrete-time stochastic process $T = T^{x_0 i_0}$ with values in $[0, \infty)$. More precisely, $T : \mathbb{N} \cup \{0\} \times \Omega_{x_0 i_0} \to [0, \infty)$ is given by

$$T(i, x(\cdot)) = T_i(x(\cdot)), \quad T_0(x(\cdot)) = 0 \ \forall x(\cdot) \tag{8}$$

with $\mathbf{P}_{x_0 i_0}(T_i < t)$ being the probability that shift number $i$ occurs before time $t$. Note that for each realization $x = x(\cdot)$ of $X$, there corresponds a realization $t_i = t_i(x)$ of $T$ consisting of the switching times of $x$. Note that we have included 0 as the 0th shift only for technical reasons. In the sequel, we will not count this as a shift.

We note that a solution $X^{x_0 i_0}$ to $\mathcal{S}_{\epsilon h}$ exists and is unique in the sense that $\mathbf{P}_{x_0 i_0}$ exists and is unique. This follows from 1) the existence and uniqueness on $E'$ of local solutions, and 2) by the existence and uniqueness of the distribution functions $F_{ij}$, induced by $h_{ij}$, on $[0, \infty)$ which determines the switching times (see the proof of Proposition 3).

*Remark 1:* We remark that the solution $X$ is not a Markov process since for any $\Delta, \Delta' > 0$ the distribution of $x(t+\Delta)$ given $x(t)$ depends on $x(t-\Delta')$. More precisely consider an example where the point $x'$ is reached at time $t'$ following two different realizations $y = y(\cdot)$ and $z = z(\cdot)$ of $X$ and that $y \neq z$ except at time $t'$. Knowing $X_{t'-\Delta}$ for some $\Delta > 0$ we know which realization we are currently following and thus we know the likely behavior of $X$ at time $t' + \Delta$. This means that $X_{t'+\Delta}$ depends on $X_{t'-\Delta}$ when conditioned $X_{t'}$ and therefor can not be Markov.

However, define the $\Lambda$-valued continuous-time stochastic process $I = I^{x_0 i_0} : [0, \infty) \times \Omega_{x_0 i_0} \to \Lambda$ by

$$I(t, x(\cdot)) = I_t(x(\cdot)), \quad I_0(x(\cdot)) = i_0 \ \forall x(\cdot),$$

with $\mathbf{P}_{x_0 i_0}(I_t = j)$ being the probability that $x(\cdot)$ evolves according to system $j$ at time $t$. Then a Markov process $\bar{X} = \bar{X}^{x_0, t_0} : [0, \infty) \times \Omega_{x_0 i_0} \to E' \times \Lambda$ is obtained by defining

$$\bar{X}(t, x(\cdot)) = (X_t(x(\cdot)), I_t(x(\cdot))).$$

By the strict positivity assumption on the intensity function the following sufficient condition for a curve to be a realization is immediate.

*Proposition 4:* For any $\epsilon > 0$, let $x(\cdot)$ be a solution to the Cauchy problem

$$x'(t) \in f^\epsilon(x(t)) \quad \text{a.e.,} \quad x(0) = x_0, \tag{9}$$

with

$$f^\epsilon : E' \to 2^E$$
$$x \mapsto \{v \in E | v = f_i(x) \text{ for all } i \in \Lambda^n \text{ such that } x \in P_i^\epsilon\}.$$

Then $x(\cdot)$ is also a realization of a solution at $x$ to the stochastic switched system $\mathcal{S}_{\epsilon h}$ for any (strictly positive) intensity function $h$.

It is easy to construct examples of realizations which are not AC. Hence there are realizations which are not solutions to (9). These solutions are the obstruction for (9) not being a necessary condition also. Luckily the set of these realizations can be neglected as we will show. For this, we first prove the following lemma.

*Lemma 1:* Let $x(\cdot) \in \Omega_{x_0 i_0}$ be a realization of a solution to the stochastic switched system $\mathcal{S}_{\epsilon h}$

1. If $x(\cdot)$ is absolute continuous, then it is a solution to the Cauchy problem (9).
2. If $x(\cdot)$ consists of finitely many local solutions, then it is a solution to the Cauchy problem (9).
3. Any restriction of $x(\cdot)$ to a bounded interval where it consists of finitely many local solutions is a solution to the Cauchy problem (9).

*Proof:* Statement 1 follows immediately from the definition of a realization, and 3 follows immediately from 2. Hence, we need only to prove 2.

We will show that $x(\cdot)$ is AC; hence, by 1. the proof will be complete. Let $x(\cdot)$ consist of $k$ local solutions. For $i = 1, ..., k$, let $x^i : J_i \to E'$ denote the local solutions making up $x : [0, \infty) \to E'$, i.e., $x(t) = x^i(t)$ for all $t \in J_i$ and $[0, \infty) = \cup_i J_i$. Since $x^i$ is AC we may find $\delta_i > 0$ for each given $\epsilon_i > 0$ such that

$$\sum_j |x(\beta_j^i) - x(\alpha_j^i)| < \epsilon_i \tag{10}$$

for any finite family of disjoint intervals $\{[\alpha_j^i, \beta_j^i]\}$ with $[\alpha_j^i, \beta_j^i] \subset J_i$ for each $j$ and $\sum_j |\beta_j^i - \alpha_j^i| < \delta_i$. Hence, for a given $\epsilon > 0$ let $\delta = \min_i\{\delta_i\}$, where $\delta_i$ is such that (10) holds true with $\epsilon_i < \epsilon/k$. It follows that, for any finite family of disjoint intervals $\{[\alpha_l, \beta_l]\}$ with $[\alpha_l, \beta_l] \subset J$ for each $l$ and $\sum_l |\beta_l - \alpha_l| < \delta$, we have

$$\sum_l |x(\beta_l) - x(\alpha_l)| = \sum_{i,j} |x(\beta_j^i) - x(\alpha_j^i)| < \sum_i \epsilon_i < \epsilon \tag{11}$$

with $\{[\alpha_j^i, \beta_j^i]\} = J_i \cap \cup_l [\alpha_l, \beta_l]$ the restriction of $\{[\alpha_l, \beta_l]\}$ to $J_i$. This proofs AC of $x(\cdot)$ and hence completes the proof of the lemma. ■

We now show that with probability one any realizations on a finite interval is a solution to the Cauchy problem (9). Let us remark that this is enough for most purposes.

*Theorem 1:* Let $X = X^{x_0 i_0}$ be a solution to the stochastic switched system $\mathcal{S}_{\epsilon h} = (E', K, G, \epsilon, h)$ and $J \subset [0, \infty)$ denote a finite interval containing 0. Assume that $f_j \in G$ is bounded for all $j \in \Lambda$, and let $O = O(X, J)$ be the set of realizations of $X$ which consists of infinitely many local solutions when restricted to $J$. Then $O$ has probability measure zero. Hence by Lemma 1 item 2, any realizations restricted to any finite interval is a solution to the Cauchy problem (9) with probability one.

*Proof:* To ease notation let $\Omega = \Omega_{x_0 i_0}$ and $\mathbf{P} = \mathbf{P}_{x_0 i_0}$. The set $O$ can equivalently be described as the set of realizations which have infinitely many shifts in finite time, i.e.,

$$O = \{x \in \Omega \mid \lim_{i\to\infty} t_i(x) \in J\} = \{\lim_{i\to\infty} t_i(x) \in J\}, \tag{12}$$

where we recall that $\{t_i(x)\}$ is the sequence of switching times of $x$. As in (12), we leave out the notation $x \in \Omega$ for any subset of $\Omega$ in the sequel.

Let $H_i$ denote the accumulated intensity from time $t_i$, i.e.,

$$H_i(t) = \int_{t_i}^{t_i+t} h_{j(t_i)\bullet}(x(t))dt,$$

and recall [13, p. 258 (Lemma 7.4II)] that

$$\mathbf{P}(\{t_{i+1}(x) - t_i(x) < \delta\}) = \mathbf{P}(\xi_i < H_i(\delta)), \tag{13}$$

where $\xi_i$ is an exponentially distributed random variable with mean 1. Now let $p(\delta)$ denote the proposition $\forall \delta > 0\ \exists N(\delta) > 0$ and consider the inclusions

$$\begin{aligned} O &\subseteq \{t_i(x) \subset J \text{ is Cauchy}\} \\ &\subseteq \{p(\delta) \ : \ t_{i+1}(x) - t_i(x) < \delta\ \forall i \geq N(\delta)\} \\ &= \{p(\delta) \ : \ H_i(t_{i+1}(x) - t_i(x)) < H_i(\delta)\ \forall i \geq N(\delta)\} \quad (14) \\ &\subseteq \bigcap_{\delta>0} \{H_i(t_{i+1}(x) - t_i(x)) < H_i(\delta)\ \forall i \geq N(\delta)\}, \quad (15) \end{aligned}$$

where in (14) we have used that $H_i$ is strictly increasing and where $N(\delta)$ in (15) is chosen as the infimum of all $N > 0$ such that $t_{i+1}(x) - t_i(x) < \delta$ for all $i \geq N$. Hence,

$$\begin{aligned} \mathbf{P}(O) &\leq \mathbf{P}\left(\bigcap_{\delta>0} \{H_i(t_{i+1}(x) - t_i(x)) < H_i(\delta)\ \forall i \geq N(\delta)\}\right) \\ &\leq \inf_{\delta>0} \mathbf{P}\left(\{H_i(t_{i+1}(x) - t_i(x)) < H_i(\delta)\ \forall i \geq N(\delta)\}\right) \\ &= \inf_{\delta>0} \mathbf{P}\left(\{\xi_i < H_i(\delta)\ \forall i \geq N(\delta)\}\right) \quad (16) \\ &\leq \inf_{\delta>0} \prod_{i\geq N(\delta)} \mathbf{P}(\{\xi_i < H_i(\delta)\}), \quad (17) \end{aligned}$$

where (13) has been used to obtain (16), and (17) follows from the independence of the $\xi_i$'s.

Assume first that there exists an infinite set $M(\delta) \subseteq \{N(\delta), N(\delta)+1, N(\delta)+2, \cdots\}$ and a function $H(\delta) > 0$ such that $H_i(\delta) \leq H(\delta)$ for all $i \in M(\delta)$. As a consequence

$$\begin{aligned} \prod_{i\geq N(\delta)} \mathbf{P}(\xi_i < H_i(\delta)) &\leq \prod_{i\in M(\delta)} \mathbf{P}(\xi_i < H_i(\delta)) \\ &\leq \prod_{i\in M(\delta)} \mathbf{P}(\xi_i < H(\delta)) \\ &= \prod_{i\in M(\delta)} (1 - e^{-H(\delta)}) = 0. \end{aligned}$$

By (17) we conclude that a realization with a bounded infinite shift time sequence and an upper bound, $H$, on infinitely many of the accumulated intensity functions, $H_i$, has probability zero of occurring.

Conversely, assume that there exists $\delta > 0$ and no infinite subset of the family of accumulated intensities $\{H_i\}_{i \geq N(\delta)}$ which can be bounded by a single function. In particular, take the whole set $\{H_i\}_{i \geq N(\delta)}$, then

$$\begin{aligned} &\forall k > 0 \ \exists j \in \{N(\delta), N(\delta)+1, N(\delta)+2, \cdots\} \\ &\text{such that } H_j(\delta) \geq k. \end{aligned} \tag{18}$$

We will argue that this is not possible. To do so we split the argument into two cases: that there is (or is not) a lower bound on the curve length of a realization between two consecutive shift times.

If there is a lower bound, this implies $\lim_{i\to\infty} |\dot{x}(t_i)| = \infty$. However, this is impossible since $|\dot{x}(t_i)| = |f_{j(i)}(x(t_i))|$ and $f_{j(i)}$ is bounded by assumption. If there is no lower bound then by (18) we conclude that $\lim_{i\to\infty} x(t_i)$ belongs to the boundary of the $\epsilon$-neighborhood. Hence, for any constant $c > 0$ there exists $N' > 0$ such that $H_{2k}(t_{2k+1} - t_{2k}) \leq c$ (or $H_{2k-1}(t_{2k} - t_{2k-1}) \leq c$) for infinitely many $k \in \{L, L+1, \cdots\}$ with $L = \max\{N, N'\}$. But this contradicts (18), completing the proof. ■

In summary, assume that we are presented with a switched system with switching noise which can be modeled as a stochastic switched system $\mathcal{S}_{\epsilon h}$ for some $\epsilon$ and $h$. We can then approach the analysis of this system in two ways; one via the stochastic process $X$ which is useful when simulations of $\mathcal{S}_{\epsilon h}$ are required, and one via (9) if a more detailed analysis is required.

In the sequel, we focus on the first approach, i.e., how can we use $X$ to study a switched system with switching noise.

### C. Simulation algorithm

We describe how to simulate the solutions at $x_0$ (starting in system $i_0$) for a switched system $\mathcal{S}$ under switching noise. That is, we describe how to simulate the solution at $x_0$ (starting in system $i_0$) to $\mathcal{S}_{\epsilon h}$ for some $\epsilon > 0$ and (strictly positive) intensity function $h$.

1. Given initial conditions $x_0$ and $i_0$, compute the local solution at $x_0$ starting in system $i_0$ for a long period of time. As a result, we obtain a sequence of points $p_0, p_1, \ldots, p_k$ with the index corresponding to the sampling times (here it is taken to be unit sampling time for simplicity, in practice this might not be possible).
2. The shift to system $j$ is simulated as follows. The part of the solution (from 1.) which lies within the first shift region (the $\epsilon$ neighborhood) is identified, say $p_u, p_{u+1}, \ldots, p_{u+v}$ (here we assume that $p_u$ and $p_{u+v}$ are approximately at the boundary of the epsilon neighborhood). The intensity is calculated at the points $h(p_{u+l}),\ l = 1, \ldots, v$ on that part of the curve and the integral (of the intensity) is computed sequentially $\sum_{j=1}^{l} h(p_{u+1}),\ l = 1, \ldots, v$. An exponentially distributed random variable $\xi$ with mean 1 is simulated, and compared with the sequence of numbers obtained by the integral computation. The shift then occurs at that point where the integral exceeds the exponentially distributed variable, i.e., for the point $p_{u+v'}$ where $\xi \leq \sum_{j=1}^{v'} h(p_{u+j}),\ \sum_{j=1}^{v'-1} h(p_{u+j}) \leq \xi,\ 1 \leq v' \leq v$ (with the second inequality being void if $v' = 1$). The above follows Algorithm 7.4III in [13, p. 260]. The process then starts over by computing the solution (obtained from system $j$) at the shift point.

## IV. Curve intensity

In principle, the distribution of the solution holds all information needed for analyzing a system with a stochastic shifting structure, but the distribution itself is quite complex and difficult to visualize. To solve this, we construct the curve intensity function (or measure) to summarize the typical behavior of a solution, and using this, we visualize the distribution of the solution through the intensity plot.

### A. Construction of intensity measure and function

Let $X$ be the solution to $\mathcal{S}_{\epsilon h}$ at $x_0 \in E'$. For any realization $x = x(\cdot)$ and each pair $(A, I)$ with $A \subseteq E'$ and $I \subseteq [0, \infty)$, we define $Z_x(A, I)$ to be the arc length of $x(I) \cap A$. Since by construction $Z_x$ is non-negative and countable additive we obtain.

*Proposition 5:* $Z_x$ is a locally finite measure on the Borel $\sigma$-algebra on $E' \times [0, \infty)$.

Intuitively $\Omega_{x_0} \to [0, \infty);\ x \mapsto Z_x(A, I)$ is a non-negative random variable. Hence, $Z(A, I)$ has a mean which will be called the curve intensity measure and denoted $\mu(A, I)$. The term measure is justified by the following.

*Proposition 6:* $\mu$ is a measure on $E' \times [0, \infty)$.

*Proof:* To prove countable additivity let $\{B_i\}$, with $B_i = (A_i, I_i)$, be a family of disjoint sets. We then have

$$\begin{aligned} \mu(\cup_i B_i) =& \mathbf{E}(Z_x(\cup_i B_i)) = \mathbf{E}(\sum_i Z_x(B_i)) \\ =& \sum_i \mathbf{E}(Z_x(B_i)) = \sum_i \mu(B_i), \end{aligned}$$

where we have used that $Z_x$ is countably additive, and then Lebesgue's monotone convergence theorem (or Fubini's theorem). ■

Let $\lambda$ denote the Lebesgue measure on $E' \times [0, \infty)$. Using the Lebesgue-Radon-Nikodym Theorem, we may split $\mu$ into two measures $\mu_a$ and $\mu_s$ such that $\mu = \mu_a + \mu_s$ with $\mu_a$ absolutely $\lambda$-continuous, and $\mu_s$ and $\lambda$ mutually singular. Roughly speaking, these two measures have the following interpretation: $\mu_s(A, I)$ is the mean length of the deterministic part of the trajectory, while $\mu_a(A, I)$ is the mean length of the stochastic part. In most examples, the deterministic part would be the part of the trajectory until the first time the trajectory hits the $\epsilon$-neighborhood, and the stochastic part is the part of the trajectory occurring after the first shift. However, the part of the curve between the first time the $\epsilon$-neighborhood is hit and the first shift occurs is a intermediate area, where the trajectory contributes to both measures.

More importantly, also by the Lebesgue-Radon-Nikodym Theorem, we obtain a $\lambda$-density of $\mu_a$, denoted $\phi$ and given by

$$\mu_a(A, I) = \int_{(A,I)} \phi(x, t) d\lambda$$

We call $\phi$ the curve intensity function.

Heuristically, the number $\phi(x, t)$ is the mean curve length in an infinitesimal area centered at the point $(x, t)$ in the space $E' \times [0, \infty)$. By integrating out $t$, we obtain a function $\phi_I(x) = \int_I \phi(x, t)$, with $I \subseteq [0, \infty)$, describing the mean curve length in an infinitesimal area centered at the point $x$. We call $\phi_I$ the mean curve intensity function, and use this as our main tool for analyzing trajectories of $\mathcal{S}$ with switching noise.

### B. Intensity plot

We illustrate the mean behavior of a switched system with switching noise by means of the mean curve intensity function $\phi_I$. However, $\phi$ (and hence $\phi_I$) is obtained via an (non-constructive) existence result, so $\phi$ is rarely explicitly available. Instead, we approximate $\phi_I$ by simulation as follows.

- Fix time interval $I \subseteq [0, \infty)$ and a rectangle $A \subseteq E'$.
- Simulate $n$ realizations $x_1, \ldots, x_n$.
- Divide $A$ into small rectangles $A_i$
- Approximate $\phi_I$ on $A_i$ by $\mathbf{E}[Z(A_i, I)] \approx \frac{1}{n} \sum_j Z_{x_j}(A_i, I)$.

Note that $Z_{x_j}(A_i, I)$ is approximated by the number of points falling in $A_i$ (since solutions are approximated by points).

## V. Example

In this section, we illustrate how the above developed theory can be used. Let us assume that we are given a switched system $\mathcal{S} = (E', K, \{f_i\})$ where $E' = \mathbb{R}^2$, $K^2 = \{P_i, P_2, P_3, P_4, P_5, P_6\}$ with $P_i$ as depicted in Figure 3, and the local dynamics, see Figure 4(a), given by

$$f_1 = f_4 = \begin{bmatrix} 3 & 0 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 5 & 1 \end{bmatrix}^{-1}$$

$$f_2 = f_5 = \begin{bmatrix} 3 & 0 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -3 & 0 \\ 5 & 1 \end{bmatrix}^{-1}$$

$$f_3(x) = f_6(x) = -x.$$

It is straight forward to see that 0 is an equilibrium point for the global dynamics (1); note however that the relaxed solutions along the $x_2$-axis have unstable behavior (i.e., 0 is a weakly stable equilibrium, [14], for the differential inclusion (3)). This fact can have consequences in applications as we now illustrate. For this purpose assume that the system has to be confined to some (safety) region, say a ball as in Figure 3, with a high probability. This yields a region of possible initial conditions, i.e., initial conditions yielding solutions belonging to the ball. In Figure 3(b), this region is illustrated as the shaded and crossed area. However, if there is switching noise, we need to make this region smaller, as shown by the shaded area in Figure 3(b). To illustrate this, assume that the radius of the (safety) ball is 5, and that the disturbances can be modeled by the stochastic switched system $\mathcal{S}_{\epsilon h} = (\mathcal{S}, \epsilon, h)$ with $\epsilon = 0.1$ and the intensity function $h$ corresponding to the uniform density function. Let us concentrate on a region around the $x_2$-axis. Here, the region of possible initial conditions is approximately bounded by the (two) local solutions starting at $(0, 1.53)$, see Figure 4(a). If we now simulate some realizations of the solution $X^{(0,1)1}$ of $\mathcal{S}_{\epsilon h}$, see Figure 4(b), we see that we need to choose a smaller region of acceptable initial conditions in order to obtain robustness of the state constraints. Here, intensity plots can help to guide in the choice of this region of acceptable initial conditions. In Figure 5, the intensity plot based on 2000 simulations of $X^{(0,1)1}$ is presented. This plot gives a measure of how frequently solutions violate the safety region. E.g. if only very few violations are allowed, then the intensity plot suggests that the safety region should have a boundary lower than the two local solutions starting at $(0, 1)$, as the intensity around the point $(0, 1.53)$ is still notable.

## References


[1] M. L. Bujorianu and J. Lygeros, "Stability and robustness for hybrid systems," in *43rd IEEE Conference on Decision and Control*, 2004.

[2] C. G. Cassandras and J. Lygeros, Eds., *Stochastic hybrid systems*, ser. Automation and Control Engineering. Boca Raton, FL: CRC Press, 2007.

[3] J. Lygeros and M. Prandini, "Stochastic hybrid systems: a powerful framework for complex, large scale applications," *Eur. J. Control*, vol. 16, no. 6, pp. 583–594, 2010. [Online]. Available: http://dx.doi.org/10.3166/ejc.16.583-594

[4] M. L. Bujorianu, "Extended stochastic hybrid systems and their reachability problem," in *HSCC*, 2004, pp. 234–249.

[5] J. Hu, J. Lygeros, and S. Sastry, "Towards a theory of stochastic hybrid systems," in *Hybrid systems: computation and control*, ser. Lecture Notes in Comput. Sci. Berlin: Springer, 2000, vol. 1790, pp. 160–173.

[6] J. P. Hespanha, "Modelling and analysis of stochastic hybrid systems," *IEE Proc. Control Theory Appl.*, vol. 153, no. 5, pp. 520–535, 2006. [Online]. Available: http://dx.doi.org/10.1049/ip-cta:20050088

[7] J. Leth and R. Wisniewski, "On formalism and stability of switched systems," *Control Theory and Applications*, vol. 10, no. 2, 2012.

[8] D. Liberzon, *Switching in systems and control*, ser. Systems & Control: Foundations & Applications. Boston, MA: Birkhäuser Boston Inc., 2003.

[9] M. Johansson and A. Rantzer, "Computation of piecewise quadratic Lyapunov functions for hybrid systems," *IEEE Trans. Automat. Control*, vol. 43, no. 4, pp. 555–559, 1998.

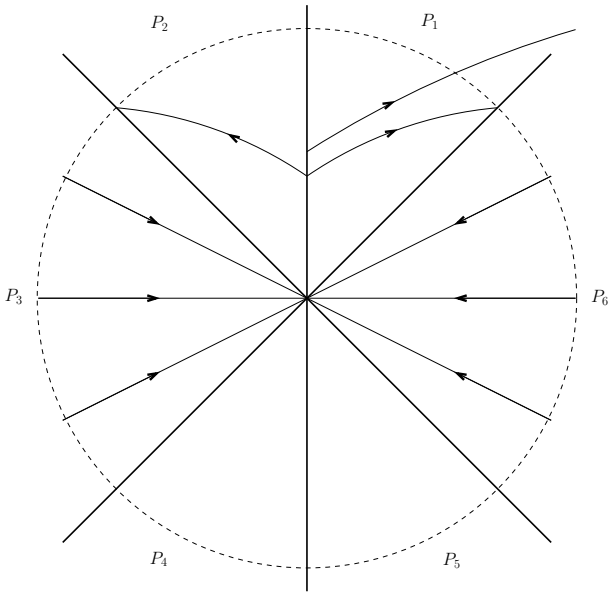
[10] A. Rantzer and M. Johansson, "Piecewise linear quadratic optimal control," *IEEE Trans. Automat. Control*, vol. 45, no. 4, pp. 629–637, 2000.

[11] J.-P. Aubin and A. Cellina, *Differential inclusions*, ser. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Berlin: Springer-Verlag, 1984, vol. 264, set-valued maps and viability theory.

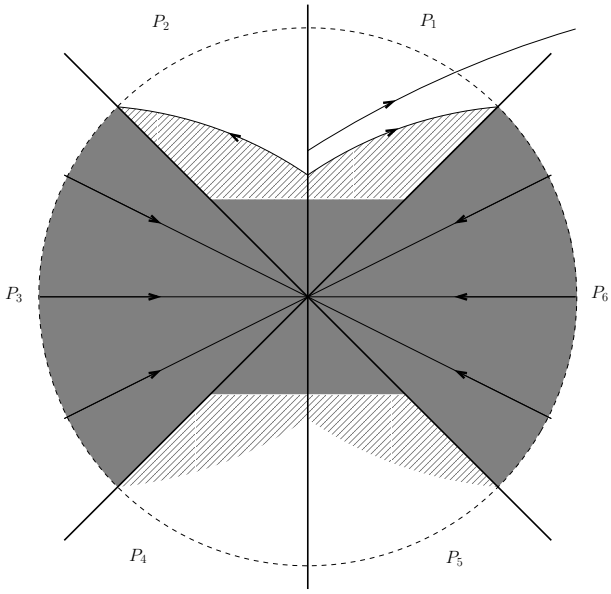
[12] D. R. Cox, *Renewal theory*. London: Methuen & Co. Ltd., 1962.

[13] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes. Vol. I*, 2nd ed., ser. Probability and its Applications (New York). New York: Springer-Verlag, 2003, elementary theory and methods.

[14] G. V. Smirnov, *Introduction to the theory of differential inclusions*, ser. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2002, vol. 41.

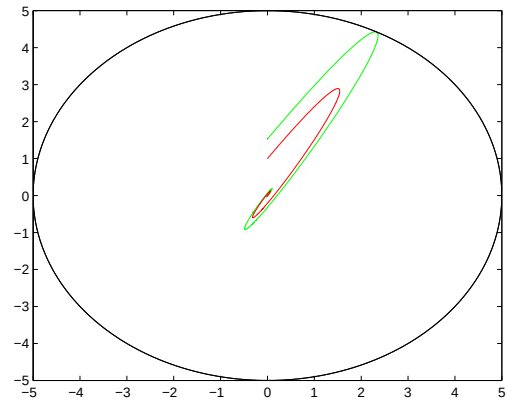



(a)

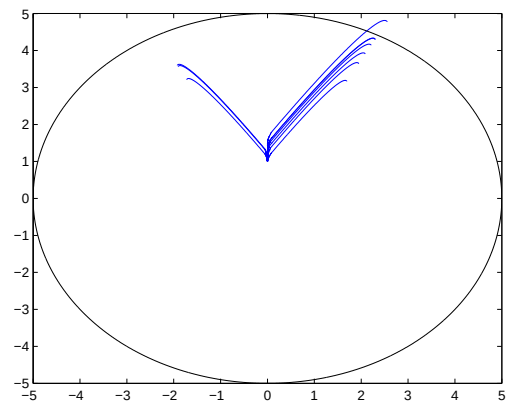



(b)

Fig. 3. Illustration of the switched system $\mathcal{S}$. (a) The three thick lines (two diagonal lines and the $x_2$-axis) indicate the switching lines, the lines with arrows indicate (local) trajectories of the system, and the stippled circle indicate the boundary of the region in which the system has to stay. In (b) the shaded and crossed area together indicate the region of possible initial conditions while the shaded area alone indicate the region of acceptable initial conditions.




(a)




(b)

Fig. 4. (a) Two local solutions, one at $(0, 1)$ and one at $(0, 1.53)$, of the local system governed by $f_1$. (b) 10 realizations of $X^{(0,1)1}$.

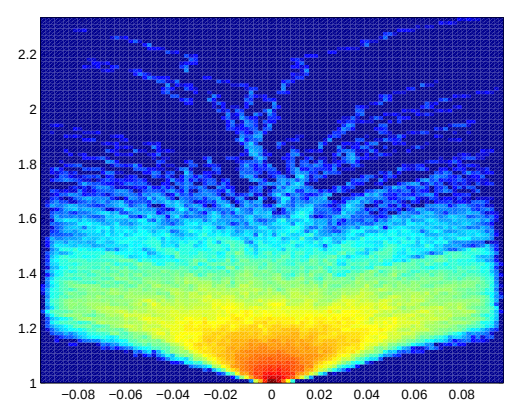



Fig. 5. Intensity plot based on 2000 realizations.