p-Safe Analysis of Stochastic Hybrid Processes

Rafael Wisniewski[1], Manuela L. Bujorianu[2], Christoffer Sloth [3]

***Abstract*—We develop a method for determining whether a stochastic system is safe, i.e., whether its trajectories reach unsafe states. Specifically, we define and solve a probabilistic safety problem for Markov processes. Based on the knowledge of the extended generator, we are able to develop an evolution equation, as a system of integral equations, describing the connection between unsafe and initial states. Subsequently, using the moment method, we approximate the infinite-dimensional optimisation problem searching for the largest set of safe states by a finite dimensional polynomial optimisation problem.**

In particular, we address the above safety problem to a special class of stochastic hybrid processes, namely piecewise-deterministic Markov processes. These are characterized by deterministic dynamics and stochastic jumps, where both the time and the destination of the jumps are stochastic. In addition, the jumps can be both spontaneous (in the style of a Poisson process) or forced (governed by guards). In this case, the extended generator of this process and its corresponding martingale problem turn out to be defined on a rather restricted domain. To circumvent this difficulty, we bring the generalized differential formula of this process into the evolution equation and, subsequently, formulate a polynomial optimisation.

***Index Terms*—Safety verification, martingale problem, moment method, stochastic hybrid systems, Markov processes, optimisation.**

I. Introduction

We develop a method for safety verification of stochastic hybrid systems. Safety verification plays an important role as the mean of examining whether a system works as intended. Routinely, a system is said to be safe if it does not violate any system constraints. Safety, defined in this way, has been studied using the concept of barrier certificates [1] and has been applied for e.g. emergency shutdown of a wind turbine [2]. The original inquiry has been later modified to cope with the design of a controller that makes the closed-loop system safe [3]. Subsequently, Romdlony and Jayawardhana [4] have noticed that the combination of the concept of barrier certificate with Lyapunov stability theory can be used to control a system with constraints, and thereby the feedback system is safe and asymptotically stable by design.

Nonetheless, in probabilistic settings, the notion of safety has first to be carefully defined. On one hand, it has to capture the nature of the technical inquiry, on the other hand, it has to be mathematically sound; and, in the best instance, it has to be numerically tractable. Indeed, the definition of probabilistic safety is the first contribution of this work. Intuitively, a system is said to be safe if it violates system constraints with a probability of at most p - an a priori assigned probability. In engineering practice, to assign the number p, one takes into account the risk and the price of a machine/infrastructure to be examined for safety, as risk = $p \times$ price.

We propose a classification of the initial states. This classification is necessary since the reachability probabilities do not capture information about the initial states that have a bigger importance and hence bigger "weight" in the computation. This formulation translates into the definitions of strong and weak p-safety. The strong p-safety is pertinent when the safety is examined for each initial point in the state space; whereas, weak p-safety makes sense when the examination of safety takes into account the distribution μ_0 of the initial states. We will not dwell on them right now, but in the next section, we will give an informal introduction and in Section IV, we will define the concept of strong and weak p-safety.

The next step, in Section V, is to develop an algorithm for the computation of the probabilistic safety - the computation of the maximal safety sets. Our choice is to lean upon the generalised moment method [5] - a particular instance of a linear infinite optimisation problem. Importantly, there are available software tools for the generalised moment method [6]. We are motivated by [7] and [8], which used the concept of an occupation measure to study the region of attraction for dynamical systems. The time evolution of the occupation measure was formulated, and subsequently translated to the generalized moment method. Nonetheless, in a stochastic setting the path of getting to the occupation measure is very different. It starts from a definition of an extended generator and Dynkin's formula. After introducing the hitting measure, which is the measure defined using the first hitting time of a specified set (usually unsafe), we formulate the adjoint equation, which relates the initial measure, the occupation measure and the hitting measure. This equation is linear but infinite dimensional. For a triplet of measures (μ, μ_0, κ), the problem of finding a process whose occupation measure is μ, the initial measure is μ_0, and the hitting measure is κ is a version of the martingale problem. It has been studied in [9] and [10]. Under some technical assumptions discussed in Section V, the martingale problem has a unique solution [11]. Subsequently, we use this result to formulate in Section VI an optimisation scheme for computing probabilistic safety by means of the generalised moment method. The results of this work apply immediately to diffusion processes and jump diffusion processes. In Section VIII, we apply the results to a specific stochastic hybrid system - the Davis' piecewise

[1]Department of Electronic Systems, Automation and Control, Aalborg University, 9220 Aalborg, Denmark raf@es.aau.dk, cs@es.aau.dk
[2]Maritime Safety Research Center, Department of Naval Architecture, Ocean & Marine Engineering, University of Strathclyde, Scotland, UK, luminita.bujorianu@strath.ac.uk
[3]University of Southern Denmark, SDU Robotics, Odense, Denmark, chsl@mmmi.sdu.dk
This work was partially supported by the Danish Council for Independent Research within the Programme Technology and Production Sciences, under the CodeMe (Computer Aided Design Methods for Industrial Automation) research project.

deterministic Markov process (PDMP) [12]. This consists of deterministic dynamical systems that alternate with random discrete transitions. Randomness of the discrete transitions (jumps) is characterised by stochastic time of jumps, and stochastic jump-destinations. We face a challenge that the domain of the extended generator of PDMP does not satisfy one of the conditions for the existence of a solution of the martingale problem in [11]. Therefore, instead of working with the extended generator, we apply the PDMP differential formula.

The problem addressed in this paper can be seen as dual to the well-known viability problem for control systems [13]. For each initial condition, viability aims to calculate the control strategy that keeps the system in a safe set. In this work, a control strategy is given, and the aim is to determine if the initial states are safe (in the probabilistic sense). We do not assume that the system is ergodic, or has any form of asymptotic stability nor there is a finite invariant measure. As a consequence, the choice of the initial states indeed matters for safety.

Alternative methods for the state constrained reachability have been proposed. Barrier certificates have been used to assess the worst case stochastic safety in [14]. Dynamic programming has been used to compute the reach-avoid probability. This approach has been developed for (deterministic) hybrid system in [15], and for discrete-time stochastic hybrid systems in [16]. The reach-avoid problem for diffusion processes with control have been characterized in [17]. Specifically, the work proposes a method for computing the initial states for which there exists a control policy such that the process reaches a goal set prior to entering a forbidden set. The approach in [17] is developed for diffusion processes, which have continuous realizations; whereas, our method is developed for hybrid systems with discontinuous realisations due to both spontaneous and forced jumps. In contrast to our work utilizing probability distribution and the evolution equation, the approach in [17] is trajectory based. Specifically, the reach avoidance probability is the level set for the viscosity solution of the Hamilton-Jacobi-Bellman equation. We see the two approaches as being a duality. Indeed, they have the same root - the Dynkin's Formula. Another related work is presented in [18]. The authors propose a numerical method for computing the survivor function for the exit time from a set. They use two observations: 1) the sequence of the times and the destination of jumps define a Markov chain 2) the sojourn time between the jumps is deterministic. The resulting numerical method leans upon a quantized approximation of the underlying Markov chain. Again, as in [17], the method in [18] is trajectory based.

The paper is organised as follows. Section II uses a simple diffusion process to give an overview of the method developed in this work and to see the difference with existing methods [16], [17]. Notations used throughout the article are introduced in Section III. Probabilistic safety concepts are defined in Section IV. The martingale problem and the adjoint equation are derived in Section V. The verification is formulated as an optimisation problem in Section VI. The findings of the paper are applied for safety verification of a PDMP. For consistency, the PDMP and its extended generator are presented in Section VII. In Section VIII, we derive the adjoint equation for PDMP. In Section IX, we provide an illustrating numerical example.

II. Overall Approach

We will briefly and informally describe the contents of the paper. The aim is to provide intuition and give the guidelines of how to use the results presented in the main body of the paper. This short section is thought as a remedy of the formal character of the paper. To make it more concrete, we exemplify the ideas and the results for a not-yet-hybrid system, a diffusion process on $\mathbb{R}^n$, of the form

$$\dot{x} = f(x) + \sigma(x)w, \tag{1}$$

where w is white noise. Following the convention from stochastic calculus, we write (1) as the following stochastic differential equation

$$dX_t = f(X_t)dt + \sigma(X_t)dB_t, \tag{2}$$

where (B_t) is the Brownian motion with values in an Euclidean space $\mathbb{R}^m$. We have used the parenthesis to indicate that a process is a sequence of random variables.

We want to study how the process (X_t) develops in a set $\mathcal{Y}$, which is a subset of $\mathbb{R}^n$. We suppose that $\mathcal{Y}$ is basic semi-algebraic, i.e., a set defined by a finite number of polynomials with real coefficients $g_i(X) \in \mathbb{R}[X]$, $i \in \{1, \dots, N\}$

$$\mathcal{Y} = \{x \in \mathbb{R}^n |\ g_i(x) \geq 0, i \in \{1, \dots, N\}\}.$$

We consider a subset $S \subset \mathbb{R}^n$ of allowable states and a subset of forbidden states $U \subset S$. Also these sets are assumed to be basic semi-algebraic. We ask the question what is the probability that (X_t) hits U before leaving S. We denote by ζ_S the first time of leaving S, and by τ_U the first time of hitting U, we re-formulate the above question as follows: What is the probability that the time of hitting U is less than the time of leaving S? Also we do want to restrict the study to some finite time horizon T, where the question of safety is still relevant. We denote this probability by $\mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T]$ with $y = X_0$. In the applications, the set U can be regarded as a set of forbidden states; whereas, the compliment of S as a goal set. Hence, ζ_S is the time when the goal is reached. On the other hand, the examples of the time T are the life-time of a system and the time to a new task.

For a number $p \in [0, 1]$, we say that a point $y \in S$ is p-safe if

$$\mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] \leq p.$$

The collection of all p-safe points will be called the largest strongly p-safe set. To find the largest strongly p-safe set, we determine the probability that the realisations of (2) hit U before leaving S and before the time passes T, for each point y. A naive method of computing $\mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T]$ is to simulate the realisations (2) by the Euler schemes [19]. In this work, we strive to represent $\mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T]$ in terms of a solution to a conic optimisation, which will be discussed later in this section.

The concept of strong p-safety is indeed strong, as the probability of the initial value being y is 1, as all the realisations of (X_t) start at y. To weaken this premise, we formulate the definition of weak p-safety. With a given initial distribution μ_0 of the start-points, we say that a set A is weakly p-safe if

$$\int_A \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] d\mu_0(y) \leq p.$$

In other words, to check whether A is weakly p-safe, first we make an 'experiment' of picking an initial value y with distribution μ_0, then we execute a realisation of the diffusion process (2) and check if it hits U without leaving S in time less than T with the probability no greater than p.

Fig. 1 illustrates the concept of weak and strong p-safety. The horizontal axis symbolizes the set $\mathcal{Y}$ with the subset U of forbidden states in red (solid line under horizontal axis). The probability $P(y) \equiv \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T]$ of reaching the forbidden states is assigned to each point $y \in \mathcal{Y}$. Specifically, it is 1 for the points in U. Each point y with $P(y) \leq p$ is p-safe, and the collection of these points comprises the strong p-safe set, which is indicated by green (dashed line under the horizontal axis). If we suppose that the initial distribution μ_0 on $\mathcal{Y}$ is uniform and the volume under the graph corresponding to the blue set (dotted line under horizontal axis) is no greater than p, then the blue set is weakly p-safe. In other words, a weak p-safe set, in this example, is the collection of points whose probability of reaching U is, in average, below p.

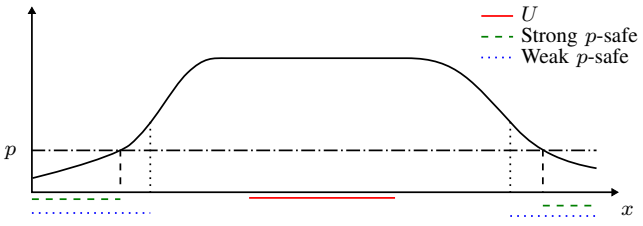

Fig. 1. Illustration of probability of reaching the forbidden states (set marked with solid red line). The strong p-safe set is highlighted with a green dashed line and the weak p-safe set is marked with a blue dotted line.

To compute p-safe sets, we propose to apply the generalised moment method [20]. In a nut-shell, the generalised moment method is a numerical scheme for solving the following optimisation over the finite Borel measures

$$\sup_{\mu} \int_{\mathcal{Y}} f d\mu \tag{3}$$

subject to

$$\int_{\mathcal{Y}} h_j d\mu \leq \gamma_j, \quad j \in \{1 \dots L\},$$

where $\mathcal{Y}$ is again a basic semi-algebraic set, f and h_j are polynomial functions on $\mathcal{Y}$, and γ_j is a real number. The solution to the above problem consists of a hierarchy of semi-definite relaxations, often referred to as the Lasserre hierarchy. The associated sequence of optimal values of the relaxed problems converges to the solution of (3) [20, Chaper 4]. This convergence has been shown to be no worse than $O(1/\sqrt{r})$, where $2r$ is the size of the semi-definite matrices involved in the relaxation [21]. Since the proposed method is based on Lasserre's hierarchy of relaxations of polynomial optimizations, the size of the optimization problem grows with the number moments, which is given by the number of monomials of a polynomial in n variables of degree d; this is $\binom{n+d}{n}$. As a consequence, the numerical method in this work applies well to dynamical systems defined on low dimensional state spaces, similarly to the method presented in [8]. Additionally, if the formulation (3) is sparse, the numerical tractability is improved [22], and larger dynamical systems can be studied.

We show in Theorem 2 that the generalised moment method can be used to characterise p-safety. Without dwelling on the details, for the initial measure μ_0, we search for a measure ν_0 on $\mathcal{Y}$ with the largest volume such that for any Borel set A, $\mu_0(A) - \nu_0(A) \geq 0$, and it satisfies the following linear equation

$$\int_{\mathcal{T}\times\mathcal{Y}} h d\kappa_1 + \int_{\mathcal{T}\times\mathcal{Y}} h d\kappa_2 - \int_{\mathcal{Y}} h(0,x) d\nu_0 = \int_{\mathcal{T}\times\mathcal{Y}} \mathcal{L}h d\mu \tag{4}$$

for any polynomial function h, where $\mathcal{T} \equiv [0, T]$. Equation (4) relates ν_0 with three yet-to-be-defined measures μ, κ_1, κ_2 with μ characterising the probability of the process (X_t) staying in S, κ_1 characterising the probability of leaving S without hitting U, and κ_2 characterising the probability of hitting the unsafe set U, which has to be no greater than p. For the diffusion process (2), the infinitesimal generator $\mathcal{L}$ is given as follows: For any differentiable function $h : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$

$$\mathcal{L}h = \frac{\partial h}{\partial t} + \frac{\partial h}{\partial x} f + \frac{1}{2} \mathrm{tr}\left(\sigma\sigma^T D^2 h\right),$$

where $\mathrm{tr}()$ stands for the trace, $\frac{\partial h}{\partial x} f = \sum \frac{\partial h}{\partial x_i} f_i$ and $D^2 h = [\frac{\partial^2 h}{\partial x_i \partial x_j}]$ is the Hessian of $h(t, \cdot)$.

Equipped with the measure ν_0 computed in (4) and the initial measure μ_0, the largest set A such that

$$(\mu_0 - \nu_0)(A) = 0,$$

is a weakly p-safe set and contains the largest strongly p-safe set. Alternatively, there is a density ρ, a nonnegative measurable function on $[0, T] \times \mathcal{Y}$ such that $d\nu_0 = \rho d\mu_0$. The set of all points x with the density $\rho(x) = 1$ is weakly p-safe. Furthermore, this set contains all p-safe points. This result can be seen as the reminiscence of a barrier certificate [14], whose value is 1 on the safe states.

This general approach works not only for diffusion processes but also for diffusion jump processes. Nonetheless, it turns out to be insufficient when dealing with PDMPs. Specifically, the Feller property is lost due to forced jumps [12, Sec. 27], and the domain of the extended generator comprises of functions with an extra condition at the boundary. This is the reason why we use the PDMP differential formula and modify (4). The details will be given in Section VIII.

In the following, we will not consider explicitly the control. Notwithstanding, we remark that by combining the findings in this work with the results in [8], the occupation measure μ in (4) can be defined on the Cartesian product $\mathcal{Y} \times \mathcal{U}$, where the control $u(t) \in \mathcal{U}$. Consequently, it is possible to characterize

the set of points for which there is a control such that the probability of the controlled process X_t leaves S before hitting U is not greater than p.

III. Notations and Definitions

For a predicate $Q(x)$ of a variable x, we will use the notation $[Q(x)]$ to denote $\{x \in A|\ Q(x)\}$ if the set A is implicitly known. For a probability $\mathbb{P}$, we write $\mathbb{P}[Q(x)] \equiv \mathbb{P}([Q(x)])$. For two functions g and h, $(g \wedge h)(x) \equiv \min\{g(x), h(x)\}$.

For a smooth function g on $\mathbb{R}^n$ and a smooth vector field f on $\mathbb{R}^n$, $L_f g$ denotes the directional derivative of g along f. The Borel sigma-algebra on a topological space $\mathcal{Y}$ is denoted by $\mathcal{B}(\mathcal{Y})$. For $A \in \mathcal{B}(\mathcal{Y})$, I_A denotes the indicator function of A. For a measure μ on a measurable space $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$, $\mathrm{supp}(\mu)$ denotes the support of μ, i.e., the largest (closed) set C such that every open set V with $V \cap C \neq \emptyset$ has positive measure, $\mu(V \cap C) > 0$. The positive cone of finite Borel measures on $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ is denoted by $\mathcal{M}_+(\mathcal{Y})$. For $\mu \in \mathcal{M}_+(\mathcal{Y})$, and $A \in \mathcal{B}(\mathcal{Y})$, we define the measure $\mu|_A(B) = \mu(A \cap B)$ for any $B \in \mathcal{B}(\mathcal{Y})$.

The complement of a set A is denoted by A^c, its closure by $\mathrm{cl}(A)$, and the boundary of A by ∂A. $\mathcal{P}(A)$ denotes the power set of A. The first time of hitting a set A by a process (X_t) is denoted by $\tau_A \equiv \inf\{t \in \mathbb{R}_{\geq 0}|\ X_t \in A\}$, and the exit time from A is defined by $\zeta_A \equiv \tau_{A^c}$. We denote the Dirac measure at y, a measure that has a unit point mass at $\{y\}$, by δ_y. The Lebesgue measure on $\mathbb{R}^n$ is denoted by λ.

We use the following notation, $B(\mathcal{Y})$ is the space of bounded measurable functions on $\mathcal{Y}$, $C(\mathcal{Y})$ is the space of continuous functions on $\mathcal{Y}$, $C_b(\mathcal{Y})$ is the space of bounded, continuous functions on $\mathcal{Y}$, and $A_b(\mathcal{Y})$ is the space of bounded, absolutely continuous functions on $\mathcal{Y}$.

We consider a process $(X_t, \mathbb{P}^y)$ on a probability space $(\Omega, \mathcal{F})$ with values in a Polish space $\mathcal{Y}$, adapted to a filtration $(\mathcal{F}_t)$. That is $X_t : \Omega \to \mathcal{Y}$ is $\mathcal{F}_t$ measurable. $(\mathbb{P}^y(A))_{y \in \mathcal{Y}}$ is the family of probability measures on $(\Omega, \mathcal{F})$ with $\mathbb{P}^y[X_0 = y] = 1$. Furthermore, we suppose that the realisations $(X_t(\omega))$ for each ω are cádlag (right continuous with left limits), and (X_t) satisfies the strong Markov property, i.e., for every $y \in \mathcal{Y}$ and every finite stopping time s, $\mathbb{P}^y[X_{s+t} \in A|\mathcal{F}_s] = \mathbb{P}^{X_s}[X_t \in A]$, $\mathbb{P}^y$-a.s. ($\mathbb{P}^y$-almost surely). To $(\mathbb{P}^y)$, we associate a transition semigroup (P_t) with $P_t(y, A) = \mathbb{P}^y[X_t \in A]$.

For a probability measure μ, we will use the notation $\mathbb{P}^\mu(B) = \int_{\mathcal{Y}} \mathbb{P}^y(B)\mu(dy)$. In other words, $\mathbb{P}^\mu[X_t \in A]$ is the probability that X_t belongs to A provided that the distribution of X_0 is μ. For a function f on $\mathcal{Y}$ and a measure μ on $\mathcal{Y}$, we define the pairing

$$\langle f, \mu \rangle = \int_{\mathcal{Y}} f(y)\mu(dy),$$

whenever the integral on the right hand side exists.

A collection of functions $\mathcal{P}$ on the space $\mathcal{Y}$ is said to separate points if for every $x, y \in \mathcal{Y}$ with $x \neq y$, there exists $h \in \mathcal{P}$ such that $h(x) \neq h(y)$.

IV. p-Safety

In this section, we define two variants of p-safety: strong and weak. In short, a set is strongly p-safe if for any point y in the set, the probability that realisations of a process starting at y reaches the unsafe states is not greater than p. Whereas, a set is weakly p-safe if the probability of reaching the unsafe set weighted by an initial measure is not greater than p.

Let S be an open and U a closed subsets of $\mathcal{Y}$ with $U \subset S$. We refer to the set S as the state space, and a point $y \in S$ as a state.

We want to determine the probability that (X_t) reaches U at some time without leaving S. The above statement can be formalised introducing the first hitting time τ_U of the set U, and the first exit time ζ_S from S. We will study the probability that the sample paths visit U before leaving S, $\mathbb{P}^y[\tau_U < \zeta_S]$. It is natural to think that if $\mathbb{P}^y[\tau_U < \zeta_S]$ is bigger than a certain threshold p then the state y is considered unsafe. Nevertheless, we will examine safety in an arbitrary but finite time-horizon T. Hence, a state y is p-safe if the probability that the process hits U within the horizon T and before it leaves S is not greater than p.

Definition 1: A point $y \in S$ is p-safe if

$$\mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] \leq p. \tag{5}$$

A state that does not satisfy (5) is called p-unsafe.

We want to determine all p-safe states.

$$\{y \in S|\ \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] \leq p\}. \tag{6}$$

Definition 2: Let μ_0 be a probability measure on the state space with $\mathrm{supp}(\mu_0) \subset S$.

- A set $A \in \mathcal{B}(\mathcal{Y})$, $A \subset S$, is strongly p-safe with respect to μ_0 if for any measurable subset $B \subset A$,
$$\chi^{\mu_0}(B) \equiv \int_B \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T]\mu_0(dy) \leq p\mu_0(B).$$
- A is strongly p-unsafe (with respect to μ_0) if for any measurable subset $B \subset A$,
$$\chi^{\mu_0}(B) > p\mu_0(B).$$

Remark 1: Notice that in the definition of strongly p-unsafe we have used the quantifier "for any" not "there exists"; hence, being strongly p-unsafe is not the same as not being strongly p-safe. Furthermore, one is tempted to replace $\chi^{\mu_0}(B) > p\mu_0(B)$ in the definition of strong p-unsafety by $\chi^{\mu_0}(B) > p$ as $p \geq p\mu_0(B)$. Nonetheless, by shrinking B, we gradually decrease $\mu_0(B)$ to zero. Hence, using this inequality, there would be no strongly p-unsafe sets.

It follows from Definition 2 that a subset $A \subset S$ with $\mu_0(A) = 0$ is strongly p-safe. Furthermore, a subset of a strongly p-safe set is again strongly p-safe and the disjoint union of strongly p-safe sets is strongly p-safe.

The next proposition establishes a relation between Definitions 1 and 2. It says that any strongly p-safe set consists of p-safe points.

Proposition 1: A subset $A \subset S$ is a strongly p-safe if and only if for any $y \in A$ (μ_0-a.s.), $P^y[\tau_U < \zeta_S, \tau_U < T] \leq p$.

Proof: For necessity, for any $B \subset A$, let

$$B' = \{y \in B |\ \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] \leq p\}.$$

By the premises of the proposition, $\mu_0(B \setminus B') = 0$. Hence,

$$\chi^{\mu_0}(B) = \int_{B'} \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T]\mu_0(dy) \leq p\mu_0(B).$$

We will prove the sufficiency by contradiction. To this end, suppose B is a subset of

$$\{y \in A |\ \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] > p\} \subset A$$

and $\mu_0(B) > 0$. As B is a subset of the strongly p-safe set A, it is also strongly p-safe. Hence, $\chi^{\mu_0}(B) \leq p\mu_0(B)$. But

$$\chi^{\mu_0}(B) = \int_B \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T]\mu_0(dy) > p\int_B \mu_0(dy),$$

which is a contradiction. ∎

We want to express that a set is larger than another one and at the same time to disregard points of zero μ_0 measure. Therefore, we define an equivalence relation $\simeq$ on $\mathcal{B}(\mathcal{Y})$ by $A \simeq B$ if and only if $\mu_0(A) = \mu_0(B)$. Subsequently, we define a partial order $\preceq$ on $\mathcal{B}(\mathcal{Y})/\simeq$

$$[A] \preceq [B] \iff \mu_0(A) \leq \mu_0(B), \tag{7}$$

where $[A]$ is the equivalence class containing A.

In the rest of the paper, when referring to the largest set, we mean a largest set with respect to $\preceq$ (if it exists). Specifically, A is a largest strongly p-safe set with respect to $\preceq$, if whenever there exists a strongly p-safe set B and $[A] \preceq [B]$ then $\mu_0(A) = \mu_0(B)$. That is, A and B belong to the same equivalence class, $[A] = [B]$.

We observe that the complement of a largest strongly p-safe set is strongly p-unsafe. In the next proposition, we show that a largest p-safe set and a largest p-unsafe set exist.

Proposition 2: The set

$$\{y \in S |\ \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] \leq p\}$$

is the largest strongly p-safe, and the set

$$\{y \in S |\ \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] > p\}$$

is largest strongly p-unsafe. Furthermore, the two sets are measurable (belong to $\mathcal{B}(\mathcal{Y})$).

Proof: The first statement of this proposition follows from Proposition 1. For the second statement, recall that $\mathcal{Y}$ is a Polish space, hence by Theorem 8.36 in [23], $y \mapsto Q(y) \equiv P^y[\tau_U < \zeta_S, \tau_U < T]$ is a measurable function. Therefore, the sets $Q^{-1}([0,p])$ and $Q^{-1}((p,1])$ are measurable. ∎

The last proposition was an existence result; whereas, the following proposition is an uniqueness result.

Proposition 3: The largest strongly p-safe set is unique up to measure μ_0.

Proof: Suppose that A and A' are both largest strongly p-safe sets such that $\mu_0(A \setminus A' \cup A' \setminus A) \neq 0$. Then $B = A \cup A'$ is strongly p-safe, $A \subset B$, and $\mu_0(B \setminus A) = 0$. This contradicts with the hypothesis that A is the largest strongly p-safe set. ∎

Another concept of safety discussed in this paper is weak p-safety. It seems reasonable to weight the unsafe points with the measure μ_0. To illustrate, if the initial measure of the set A comprising of unsafe points is particularly small, we might want to regard A as safe. The next definition captures this concept.

Definition 3: Let μ_0 be a probability measure on the state space S, i.e., $\text{supp}(\mu_0) \subset S$. We will say that a set $A \in \mathcal{B}(\mathcal{Y})$, $A \subset S$, is weakly p-safe (with respect to μ_0) if

$$\chi^{\mu_0}(A) \leq p.$$

We say that A is weakly p-unsafe if

$$\chi^{\mu_0}(A) > p.$$

Notice that any subset of a weakly p-safe set is again weakly p-safe, as χ^{μ_0} is a measure. Furthermore, any strongly p-safe set is weakly p-safe as

$$\chi^{\mu_0}(B) \leq p\mu_0(B) \leq p.$$

The next example shows that a weakly p-safe set can contain a strongly p-unsafe set; and vice versa a weakly p-unsafe set can contain a strongly p-safe set.

Example 1: Suppose that the initial measure μ_0 corresponds to the uniform distribution on a compact state space S. Let $p = 0.6$, and $\{A_1, A_2, A_3\}$ be a disjoint partition of S such that

1) $\mu_0(A_1) = 0.1$, $\mu_0(A_2) = 0.2$, and $\mu_0(A_3) = 0.7$;
2) for all $y \in A_1$, $P^y[\tau_U < \zeta_S] = 0.1$, and for all $y \in A_2 \cup A_3$, $P^y[\tau_U < \zeta_S] = 0.9$.

The set A_1 is obviously strongly p-safe; whereas, the sets A_2 and A_3 are strongly p-unsafe. The set $A \equiv A_1 \cup A_2$ is weakly p-safe, as $\chi^{\mu_0}(A) = 0.181$. Notice also that $A' \equiv A_1 \cup A_3$ is weakly p-unsafe, $\chi^{\mu_0}(A') = 0.631$.

Any weakly p-safe set A can be uniquely partitioned into a strongly p-safe and a strongly p-unsafe disjoint sets B and C, $A = B \cup C$, since a point is either p-safe or p-unsafe. Motivated by the above example, we ask the following question. Suppose a set A is weakly p-safe. We regard a partitioning of A into two subsets B and C such that B is strongly p-safe, and C is weakly p-safe. How small can C be? Specifically, C is greater (with respect to the inclusion) than the set of all p-unsafe points, and at most, it is equal to A. To answer this question, we formulate the following proposition.

Proposition 4: Let $A \subset \mathcal{Y}$, and let $B \subset A$ be strongly p-safe. Suppose that there is an $\epsilon \geq 0$ such that

$$\begin{aligned} C &\equiv A \setminus B \\ &= \{y \in \mathcal{Y} |\ p < \mathbb{P}^y[\tau_U < \zeta_S, \tau_U < T] \leq p(1+\epsilon),\} \end{aligned}$$

and

$$\epsilon\mu_0(C) \leq \mu_0(A^c).$$

Then A is weakly p-safe.

The number ϵ in Proposition 4 can be viewed as the degree to which C is unsafe. Specifically, Proposition 4 states that the union of a strongly p-safe set B and a strongly p-unsafe set C is weakly p-safe provided the measure μ_0 of C is small or the unsafety degree ϵ of C is small.

Proof:

$$\chi^{\mu_0}(A) = \chi^{\mu_0}(B) + \chi^{\mu_0}(C) \leq p\mu_0(B) + p(1+\epsilon)\mu_0(C)$$
$$= p\mu_0(A) + p\epsilon\mu_0(C) \leq p\mu_0(A) + p\mu_0(A^c) = p.$$

■

Remark 2: The concept of weak p-safety allows formulating safety for a chain of events. Specifically, assuming that the distribution of the initial states is given by μ_0, the probability of visiting firstly a set U_1 in a time-horizon T_1 and subsequently hitting U_2 within a time-horizon T_2 is

$$\chi^{\mu_0'}_{(U_2,T_2)}(B) \equiv \int_B \mathbb{P}^y[\tau_{U_2} < \zeta_S, \tau_{U_2} < T_2]\mu_0'(dy),$$

where

$$\mu_0' = \chi^{\mu_0}_{(U_1,T_1)}.$$

Hence, a set B is weakly p-safe for the above chain of events if $\chi^{\mu_0'}_{(U_2,T_2)}(B) \leq p$.

The aim of the following sections is to characterise the largest strongly p-safe set, and to use conic optimisation for this purpose. To capture all strongly p-safe points, we use the measures ν_0, μ, κ_1 and κ_2, which will be formally introduced in the next section, and their dynamics specified by an integral equation in (4). Nonetheless, an integral "averages" the information; hence, an algorithm leaning on (4) will allow the computation of a weakly p-safe set. By adding additional optimisation constraints, we will be able to pick a weakly p-safe that contains the largest strongly p-safe set.

In conclusion, for a Markov family $(X_t, \mathbb{P}^y)$ adapted to the filtration $(\mathcal{F}_t)$, a number $p \in [0,1]$, and a subset $U \subset S$, we will develop an algorithm for computing a weakly p-safe set that contains the largest strongly p-safe set.

V. Time-Space Process

We recall, a real-valued process (Y_t) is a martingale with respect to $(\mathcal{F}_t)$ if for every $y \in \mathcal{Y}$, $\mathbb{E}^y[Y_t|\mathcal{F}_s] = Y_s$ for $t > s$ and supermartingale if $\mathbb{E}^y[Y_t|\mathcal{F}_s] \leq Y_s$ for $t > s$. A process (Y_t) is a local martingale if there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times (with respect to $(\mathcal{F}_t)$) such that $T_n \to \infty$ pointwise and the stopped process $(Y_t^{T_n})$

$$Y_t^{T_n} \equiv \begin{cases} Y_t & \text{if } t \leq T_n \\ Y_{T_n} & \text{if } t \geq T_n. \end{cases}$$

is a martingale.

A. Generator

We will introduce the so-called extended generator [12] of the considered Markov process (X_t) - a generalisation of the infinitesimal generator used in Section II. The reason for doing this is the need for broadening the domain of the generator when dealing with more general processes than the diffusion.

Following [12], we define $\mathcal{D}(\mathcal{L})$ as the set of measurable functions $h : (\mathcal{Y}, \mathcal{B}(\mathcal{Y})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ having the property that there is a measurable function $g : (\mathcal{Y}, \mathcal{B}(\mathcal{Y})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that the function $t \mapsto g(X_t)$ is $\mathbb{P}^y$-almost surely integrable for each $y \in Y$, and the process C_t^h given by

$$C_t^h \equiv h(X_t) - h(X_0) - \int_0^t g(X_s)ds \tag{8}$$

is a local martingale (with respect to $\mathcal{F}_t$). We write $\mathcal{L}h = g$ and call $(\mathcal{D}(\mathcal{L}), \mathcal{L})$, or even $\mathcal{L}$, an extended generator. Notice that the extended generator $\mathcal{L}$ is possibly multivalued. Nonetheless, if g_1 and g_2 are two extended generators, then $g_1(x) \neq g_2(x)$ only on a subset A where $\int_0^\infty I_A(X_t)dt = 0$ $\mathbb{P}^y$-a.s. for $y \in \mathcal{Y}$, i.e., the process (X_t) spends no time in A.

In the following, we will refer to the extended generator as the generator. The generators of many interesting processes encountered in control engineering have been characterised; herein, diffusion processes, their generalisations jump diffusion processes and switching diffusion processes, as well as piecewise-deterministic Markov processes.

We turn the above problem statement upside down: Having an operator $\mathcal{L}$ with a domain $\mathcal{D}(\mathcal{L})$, we ask the question if there is a process for which C_t^h is a local martingale. Specifically, let μ_0 be the initial distribution of the desired processes. We say that a process (X_t) is a solution of the *martingale problem* for $(\mathcal{L}, \mu_0)$ up to a (finite) stopping time τ if there is a filtration $(\mathcal{F}_t)$ such that X_0 has the distribution μ_0, (X_t) is $(\mathcal{F}_t)$-progressively measurable, and

$$h(X_t) - h(X_0) - \int_0^t \mathcal{L}h(X_s)ds$$

is a martingale for every h in the domain of $\mathcal{L}$ and $t \leq \tau$. We suppose that the martingale problem is well-posed that is any two solutions have the same finite-dimensional distribution. For conditions on the generator $\mathcal{L}$ imposing well-posed martingale problem we refer the interested reader to [24]. Specifically, the martingale problem is well-posed for multivariate point processes [25], thereby for PDMPs [26].

We consider the minimum time of hitting the forbidden set U or leaving the state space S. At the outset, we fix a time horizon $T \in \mathbb{R}_+$. We think of T as the maximum time of interest, above which the information whether the system is safe or not is no longer relevant. For instance, T might be the life-time of the system. Consequently, we define the stopping time τ by

$$\tau \equiv T \wedge \tau_U \wedge \zeta_S.$$

We observe that τ is finite. Therefore, by the optional sampling theorem, the stopped process $(C_t^h)^\tau = (C_{t\wedge\tau}^h)$ is a martingale, and $\mathbb{E}^x[C_{t\wedge\tau}^h] = \mathbb{E}^x[C_0^h] = 0$. Subsequently, we have

$$\mathbb{E}^y[h(X_{t\wedge\tau})] = \mathbb{E}^y[h(X_0)] + \mathbb{E}^y\left[\int_0^{t\wedge\tau} \mathcal{L}h(X_s)\right]ds. \tag{9}$$

Eq. (9) is instrumental to our work. In the next section, we show how to re-formulate it in terms of the occupation measure, which we subsequently compute.

Since (9) is valid for an arbitrary function $h \in \mathcal{D}(\mathcal{L})$, the indicator functions I_U and I_{S^c} (assuming that their smooth approximations belong to the domain of $\mathcal{L}$) evaluated in (9)

determine whether the process (X_t) hits the forbidden set U or leaves the state space S provided $\tau < T$. Nonetheless, we encounter the difficulty with distinguishing whether $\tau < T$ or $\tau = T$, as this information is not encoded in (9). To circumvent this obstacle, we define the time-space process (S_t, X_t), where S_t is the solution of the Cauchy problem $\frac{d}{dt}S_t = 1,\ S_0 = b$, i.e., $S_t = b+t$. In other words, we extend the state space of the process (X_t) with the time dimension. Here, the time can be seen as a (deterministic) process with infinitesimal generator $\frac{\partial}{\partial t}$, Section 2.4 in [27].

Consequently, (9) pronounces

$$C^h_{t\wedge\tau} \equiv h(t\wedge\tau, X_{t\wedge\tau}) - h(0, X_0) - \int_0^{t\wedge\tau} \hat{\mathcal{L}}h(s, X_s)ds, \tag{10}$$

with $\hat{\mathcal{L}} = \mathcal{L} + \frac{\partial}{\partial t}$, and $C^h_{t\wedge\tau}$ is a martingale for any measurable function $h : (\mathbb{R}\times\mathcal{Y}, \mathcal{B}(\mathbb{R}\times\mathcal{Y})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ in the domain of $\hat{\mathcal{L}}$. Specifically, the functions of the form $\gamma(t)h(x)$ with γ continuously differentiable and $h \in \mathcal{D}(\mathcal{L})$ are in $\mathcal{D}(\hat{\mathcal{L}})$.

We extend the family of probability measures $\mathbb{P}^y$ related to the process (X_t) to the family of probability measures $\hat{\mathbb{P}}^{(b,y)}$ related to the extended process $\hat{X}_t \equiv (S_t, X_t)$ by

$$\hat{\mathbb{P}}^{(b,y)}[(S_t, X_t) \in B\times C] = I_B(t+b)\mathbb{P}^y[X_t \in C].$$

As the consequence, the following equality holds

$$\begin{aligned}&\hat{\mathbb{E}}^{(0,y)}[h(t\wedge\tau, X_{t\wedge\tau})]\\ =&\hat{\mathbb{E}}^{(0,y)}[h(0,X_0)] + \hat{\mathbb{E}}^{(0,y)}\left[\int_0^{t\wedge\tau}\hat{\mathcal{L}}h(s,X_s)ds\right],\end{aligned} \tag{11}$$

where the expected value $\hat{\mathbb{E}}^{(0,y)}$ is calculated with respect to probability $\hat{\mathbb{P}}^{(0,y)}$. We re-interpret the stopping time

$$\tau = \tau_{[0,T)\times U} \wedge \zeta_{[0,T]\times S}. \tag{12}$$

The information whether $\tau_U < T$ can be now extracted from (11) by applying the indicator function $I_{[0,T)\times U}$ (provided that the indicator functions can be approximated by the elements of $\mathcal{D}(\hat{\mathcal{L}})$). Indeed, this will be a standing hypothesis in the remainder of this paper. The left hand side of (11) becomes the probability that (X_t), for $t \le T$, hits U before leaving S.

To formulate the p-safety of a process (X_t), we have used the measure χ^{μ_0} on $\mathcal{Y}$ in Definition 2, which for the corresponding time-space process (S_t, X_t) with an initial distribution $\hat{\mu}_0$ on $\mathbb{R}_+\times\mathcal{Y}$ becomes

$$\begin{aligned}\chi^{\hat{\mu}_0}(B\times C) &= \int_{B\times C}\hat{\mathbb{P}}^{(t,y)}[\tau_{[0,T)\times U} < \zeta_{[0,T]\times S}]\hat{\mu}_0(dt\times dy)\\ &= \int_{B\times C}\hat{\mathbb{P}}^{(t,y)}[\tau_{\mathbb{R}_+\times U} < \zeta_{[0,T]\times S}]\hat{\mu}_0(dt\times dy).\end{aligned} \tag{13}$$

We extend the transition semigroup $P_t(y, C)$ to the time-space transition semigroup $\hat{P}_t((b,y), B\times C)$ (for $B \in \mathcal{B}(\mathbb{R}_+)$, and $C \in \mathcal{B}(\mathcal{Y})$) by

$$\hat{P}_t((b,y), B\times C) \equiv \delta_{t+b}(B)P_t(y, C),\ t \ge 0,$$

where δ_c is the Dirac measure at c. In fact, $(\hat{P}_t)$ is the transition semigroup of the process (t, X_t), i.e., $\hat{P}_t((b,y), B\times C) = \mathbb{P}^{(b,y)}[(S_t, X_t) \in B\times C]$.

We define an action of $\hat{P}_t$ on a bounded measurable function $h \in B([0,T]\times\mathcal{Y})$ by $\hat{P}_t h(b,y) \equiv \hat{\mathbb{E}}^{(b,y)}(h(t, X_t))$, where

$$\begin{aligned}\hat{\mathbb{E}}^{(b,y)}(h(t,X_t)) &= \int_{\mathbb{R}_+\times\mathcal{Y}} h(u,z)\hat{P}_t((b,y), dudz)\\ &= \int_{\mathcal{Y}} h(t+b, z)P_t(y, dz).\end{aligned}$$

In particular, $\hat{P}_t((b,y), B\times C) = \hat{P}_t I_{B\times C}(b,y)$, where $\hat{P}_t$ on the left hand side is the transition semigroup giving the probability of the process (S_t, X_t) belonging to the set $B\times C$; whereas, $\hat{P}_t$ on right hand side is the actions on the indicator function $I_{B\times C}$.

As indicated, we focus on the process (X_t) up to the stopping time τ. To this end, for any $t \in \mathbb{R}_+$, we define a family of probabilistic operators/kernels (K_t) by

$$K_t h(b,y) \equiv \hat{\mathbb{E}}^{(b,y)}(h(t, X_t)I_{[t\le\tau]}),$$

and

$$K_t((b,y), B\times C) \equiv K_t I_{B\times C}(b,y).$$

Observe that if $B \subset (T,\infty)$ or $C \subset S^c$ then $K_t((b,y), B\times C) = 0$. Furthermore, if $t \le \tau$, $\mathbb{P}^{(c,y)}$-a.e., then $K_t = \hat{P}_t$. Hence, (K_t) fully characterises the process (X_t) up to the stopping time τ. That is, (K_t) determines the probability that (X_t) hits a set before "dying" - leaving the state space S or entering the unsafe set U.

Likewise $(\hat{P}_t)$, (K_t) for $t \in \mathbb{R}_+$ forms a one parameter semigroup, $K_{t+s}f(b,y) = K_t(K_s f)(b,y)$. This follows from the observation that τ in (12) is a terminal time, i.e., it satisfies $s + \tau\circ\theta_s = \tau$ $\mathbb{P}$-a.s. on $[s < \tau]$, where θ_s is the shift operator for $(\hat{X}_t)$.

B. Occupation Measures

We will formulate (11) as an equation relating three measures: 1) the occupation measure capturing the information about where the realisations of the process evolve in the state space, 2) hitting measure encapsulating the information of probability of hitting the boundary of the state space and the forbidden state, and 3) the initial measure.

For an initial probability measure $\hat{\mu}_0$ on $[0,T]\times(S\setminus U)$, we define the occupation measure at time t $\hat{\mu}_t$ on $[0,T]\times\mathcal{Y}$ by

$$\hat{\mu}_t(B\times C) = \int_{\mathbb{R}_+\times\mathcal{Y}} K_t((b,y), B\times C)\hat{\mu}_0(db\times dy).$$

In the following, we study initial probability measures of the form $\hat{\mu}_0 \equiv \delta_0\otimes\mu_0$. We are ready to define the primary objects of the investigation in this work:

1) the occupation kernel

$$\begin{aligned}\bar{\mu}((b,y), B\times C) &\equiv \int_0^\infty K_t((b,y), B\times C)dt\\ &= \hat{\mathbb{E}}^{(b,y)}\int_0^\tau I_{B\times C}(t, X_t)dt,\end{aligned}$$

and the occupation measure

$$\begin{aligned}\hat{\mu}(B\times C) &\equiv \int_{\mathbb{R}_+\times\mathcal{Y}} \bar{\mu}((b,y),B\times C)\hat{\mu}_0(db\times dy)\\ &= \int_{[0,T]\times\mathcal{Y}} \bar{\mu}((b,y),B\times C)\hat{\mu}_0(db\times dy)\\ &= \int_{[0,T]} \hat{\mu}_t(B\times C)dt \end{aligned} \tag{14}$$

with $\hat{\mu}_0 \equiv \delta_0 \otimes \mu_0$,

2) the hitting kernel

$$\bar{\kappa}((b,y),B\times C) \equiv \hat{\mathbb{E}}^{(b,y)}\left(I_{B\times C}(\tau, X_\tau)\right),$$

and the hitting measure

$$\hat{\kappa}(B\times C) \equiv \int_{\mathbb{R}_+\times\mathcal{Y}} \bar{\kappa}((b,y),B\times C)\hat{\mu}_0(db\times dy). \tag{15}$$

As the time-space process (S_t, X_t) terminates when leaving $[0,T]\times(S\setminus U)$, we observe that the support of the occupation and hitting measures

$$\begin{aligned}&\mathrm{supp}(\hat{\mu}) \subset [0,T]\times \mathrm{cl}(S\setminus U),\\ &\mathrm{supp}(\hat{\kappa}) \subset \{T\}\times \mathrm{cl}(S\setminus U)\cup[0,T]\times S^c \cup [0,T]\times U.\end{aligned}$$

Equipped with the occupation and the hitting measures, (11) becomes

$$\int_{[0,T]\times\mathcal{Y}} h(t,z)\hat{\kappa}(dt\times dz) = \int_{[0,T]\times\mathcal{Y}} h(t,z)\hat{\mu}_0(dt\times dz) \tag{16}$$

$$+\int_{[0,T]\times\mathcal{Y}} \hat{\mathcal{L}}h(t,z)\hat{\mu}(dt\times dz)$$

for any h in the domain of $\hat{\mathcal{L}}$, where $\hat{\mu}_0 = \delta_0\otimes\mu_0$.

To simplify the notation, we introduce the pairing $\langle g,\rho\rangle$ for an integrable function g on $[0,T]\times\mathcal{Y}$ and a measure ρ on $\mathcal{B}([0,T]\times\mathcal{Y})$,

$$\langle g,\rho\rangle \equiv \int_{[0,T]\times\mathcal{Y}} g(t,z)\rho(dt\times dz).$$

As a consequence (16) becomes

$$\langle h,\hat{\kappa}\rangle = \langle h,\hat{\mu}_0\rangle + \langle\hat{\mathcal{L}}h,\hat{\mu}\rangle. \tag{17}$$

Definition 4: We say that a triple

$$(\mu_1,\mu_2,\mu_3) \in \mathcal{M}_+([0,T]\times\mathcal{Y})\times\mathcal{M}_+(\mathcal{Y})\times\mathcal{M}_+([0,T]\times\mathcal{Y})$$

of measures solve the evolution equation for the generator $\hat{\mathcal{L}}$ if

$$\langle h,\mu_3\rangle = \langle h,\hat{\mu}_2\rangle + \langle\hat{\mathcal{L}}h,\hat{\mu}_1\rangle. \tag{18}$$

is satisfied for every $h\in\mathcal{D}(\hat{\mathcal{L}})$.

In the remaining part of this section, we reverse the problem and assume that there is a triple $(\hat{\mu},\hat{\mu}_0,\hat{\kappa})$ that is a solution of (17). We strive to recover the original process (X_t). To this end, we recall the definition of a martingale problem in Section V-A. So far, we have explained how to compute the occupation and hitting measures for a given process (S_t,X_t). On the other hand, Theorem 3.3 in [11] answers the question if there is a process (X_t), which is a solution of the martingale problem $(\mathcal{L},\mu_0)$ up to time $\tau_U\wedge\zeta_S\wedge T$.

Assumption 1: Let $\mathcal{L}$ be an operator that satisfies

1) $\mathcal{L}:\mathcal{D}(\mathcal{L})\subset C_b(\mathcal{Y})\to C(\mathcal{Y})$, and $\mathcal{L}\mathbf{1}=0$, where $\mathbf{1}: y\mapsto 1$;
2) The domain $\mathcal{D}(\mathcal{L})$ of $\mathcal{L}$ is closed under multiplication and separates points, see Section III;
3) The graph of $\mathcal{L}$ is separable, i.e., there exists a countable collection $(h_k)\subset\mathcal{D}(\mathcal{L})$ such that $(h,\mathcal{L}h)_{h\in\mathcal{D}(\mathcal{L})}$ is contained in the bounded, pointwise closure of the linear span of $(h_k,\mathcal{L}h_k)_{k\geq 1}$.

We shed some light on Assumption 1. Condition 1) states that all $\mathcal{L}$ is zero on constant functions. In particular, if $\mathcal{L}$ is an infinitesimal generator of the process (X_t), $\mathcal{L}\mathbf{1}=0$, due to Markov semigroup property, $\mathbb{P}_t^{(t,y)}\mathbf{1}=\mathbf{1}$. Condition 2) is a necessary condition for a martingale problem to be well-posed. For details, we refer the reader to Theorem 2.1 in [10]. Condition 3) allows to represent any function f (in the domain of $\mathcal{L}$) and $\mathcal{L}f$ as a linear combination of basis functions h_k. Explicitly, Condition 2) together with 3) ensure the existence of a stationary process [24, page 247] whenever there is a measure μ on the state space $\mathcal{Y}$ such that $\int_{\mathcal{Y}}\mathcal{L}f d\mu = 0$ for all $(f,\mathcal{L}f)\in\mathcal{D}(\mathcal{L})$. This property is used in the proof of Theorem 3.3 in [11] to construct a process whose initial, hitting, and occupation measures satisfy (17).

Theorem 1 (Theorem 3.3 in [11]): Suppose that $\mathcal{L}$ is an operator with the domain $\mathcal{D}(\mathcal{L})$ that satisfies Assumption 1. Let $\hat{\kappa}$ be the probability measure and $\hat{\mu}$ be the measure both on $\mathbb{R}_+\times\mathcal{Y}$ that satisfy

$$\langle\gamma h,\hat{\kappa}\rangle = \langle\gamma h,\delta_0\otimes\mu_0\rangle + \langle\frac{d}{dt}\gamma + Lh,\hat{\mu}\rangle,$$

for all continuously differentiable functions $\gamma:\mathbb{R}_+\to\mathbb{R}$ that vanish at infinity and $h\in\mathcal{D}(\mathcal{L})$. Then there exist a filtration $(\mathcal{F}_t)$, a process (X_t), and an $(\mathcal{F}_t)$-stopping time s such that (X_t) is a solution of the martingale problem for $(\mathcal{L},\mu_0)$ up to time s, $\hat{\kappa}$ is the hitting measure and $\hat{\mu}$ is the occupation measure for the time-space process (S_t,X_t) with the stopping time s.

The following corollary is a consequence of Theorem 1.

Corollary 1: Suppose that $\mathcal{L}$ is an operator with the domain $\mathcal{D}(\mathcal{L})$ that satisfies Assumption 1. Let $(\hat{\mu},\mu_0,\hat{\kappa})$ be a triplet of measures solving the evolution equation for $\hat{\mathcal{L}}$, where μ_0 and $\hat{\kappa}$ are probability measures, and the supports of $\hat{\mu}$, μ_0, and $\hat{\kappa}$ satisfy

$$\mathrm{supp}(\hat{\mu}) \subset [0,T]\times\mathrm{cl}(S\setminus U) \tag{19a}$$

$$\mathrm{supp}(\mu_0) \subset S\setminus U, \tag{19b}$$

$$\mathrm{supp}(\hat{\kappa}) \subset \{T\}\times\mathrm{cl}(S\setminus U)\cup[0,T]\times S^c\cup[0,T]\times U, \tag{19c}$$

and the martingale problem $(\mu_0,\mathcal{L})$ well-posed.

Then

$$\chi^{\delta_0\otimes\mu_0}(\mathbb{R}_+\times S) = \hat{\kappa}([0,T)\times U). \tag{20}$$

Proof: By Theorem 1, there is a process $(Y_t,\tilde{\mathbb{P}}^y)$ a solution of the martingale problem for $(\mathcal{L},\mu_0)$, and a stopping time s such that the hitting kernel of the time-space process

(S_t, X_t) is $\hat{\kappa}$ and the occupation measure is $\hat{\mu}$. Since the martingale problem is well-posed, we have $X_t = Y_t$, a.e., up to stopping time s. Here, a.e. means $\tilde{\mathbb{P}}^{\mu_0}$-almost everywhere.

By (19c), $s \geq \tau_U \wedge \zeta_S \wedge T$, a.e., and by (19a), $s \leq \tau_U \wedge \zeta_S \wedge T$, a.e. Hence, $s = \tau_U \wedge \zeta_S \wedge T$, a.e.

Now, (20) follows from the definition of $\chi^{\delta_0 \otimes \mu_0}(\mathbb{R}_+ \times S)$ in (13) and

$$\begin{aligned}&\hat{\mathbb{P}}^{(0,y)}[\tau_{[0,T)\times U} \wedge \zeta_{[0,T]\times S}] = \hat{\mathbb{P}}^{(0,y)}[X_\tau \in U,\ \tau \in [0,T)]\\ =&\hat{\mathbb{E}}^{(b,y)}\left(I_{[0,T)\times U}(\tau, X_\tau)\right) = \bar{\kappa}((0,y),[0,T)\times U). \end{aligned} \tag{21}$$

■

In the corollary, we have assumed well-posedness of the martingale problem. Specifically, the martingale problem is well-posed for diffusion processes, Levy processes and PDMP (for PDMP the result can be derived from the results existing for multivariate point processes [24]). In the next section, we will use Theorem 1 for the verification of p-safety.

VI. Safety Verification

We strive to solve the problem of computing a maximal p-safe set applying the generalized moment method [20]. In a nutshell, the moment method computes a solution to

$$\sup_{\mu \in \mathcal{M}_+([0,T]\times\mathcal{Y})} \langle f, \mu \rangle$$

subject to

$$\langle h_j, \mu \rangle \leq \gamma_j, \quad j \in \Gamma,$$

where Γ is a finite index set, f, h_j are integrable functions on $[0,T] \times \mathcal{Y}$ and γ_j are real numbers.

To this end, we let $\hat{\kappa} = \hat{\kappa}_1 + \hat{\kappa}_2$, where

$$\hat{\kappa}_1 \equiv \hat{\kappa}|_{\{T\}\times(S\setminus U)\cup[0,T]\times S^c} \tag{22a}$$
$$\hat{\kappa}_2 \equiv \hat{\kappa}|_{[0,T)\times U}. \tag{22b}$$

The measure $\hat{\kappa}_1$ is the probability that the process (X_t) leaves the state space before hitting the unsafe set U; whereas, $\hat{\kappa}_2$ is the probability that (X_t) hits U within the time horizon T. Specifically, by (20), for an initial distribution μ_0 on $S \setminus U$, we have

$$\hat{\kappa}_2([0,T]\times\mathcal{Y}) = \hat{\kappa}([0,T)\times U) = \chi^{\delta_0\otimes\mu_0}([0,T]\times\mathcal{Y}).$$

In the following, we use the notation $\chi^{\mu_0} \equiv \chi^{\delta_0\otimes\mu_0}$ To compute the probability $\chi^{\mu_0}([0,T]\times\mathcal{Y})$ of hitting U before leaving S in the time horizon T, we lean upon the aforementioned generalized moment method.

In the next theorem, we will show that computation of the maximal p-safe set can also be formulated as a generalized moment problem. For this purpose, we define the involved measures as follows.

Definition 5: We denote by

$$\mathcal{O} \subset \mathcal{M}_+([0,T]\times\mathcal{Y}) \times \mathcal{M}_+(\mathcal{Y}) \times \mathcal{M}_+([0,T]\times\mathcal{Y})^2,$$

the set of quadruples of measures $(\hat{\mu}, \nu_0, \hat{\kappa}_1, \hat{\kappa}_2)$ whose supports satisfy

$$\operatorname{supp}(\hat{\mu}) \subset [0,T]\times \operatorname{cl}(S\setminus U), \tag{23a}$$
$$\operatorname{supp}(\nu_0) \subset S \setminus U, \tag{23b}$$
$$\operatorname{supp}(\hat{\kappa}_1) \subset \{T\}\times\operatorname{cl}(S\setminus U)\cup[0,T]\times S^c, \tag{23c}$$
$$\operatorname{supp}(\hat{\kappa}_2) \subset [0,T]\times U. \tag{23d}$$

Theorem 2 (Characterisation of the largest strongly p-safe set): Let μ_0 be an initial probability measure on $\mathcal{Y}$ with $\operatorname{supp}(\mu_0) \subseteq S \setminus U$. Suppose that the quadruple of measures $(\hat{\mu}^*, \nu_0^*, \hat{\kappa}_1^*, \hat{\kappa}_2^*) \in \mathcal{O}$ in Definition 5 is a solution to the conic optimisation[1]

$$v^* = \sup\{\nu_0(\mathcal{Y}) |\ (\hat{\mu}, \nu_0, \hat{\kappa}_1, \hat{\kappa}_2) \in \mathcal{O}\} \tag{24}$$

such that

$$\forall h \in \mathcal{D}(\hat{\mathcal{L}}) \qquad \langle h, \hat{\kappa}_1\rangle + \langle h, \hat{\kappa}_2\rangle - \langle h, \delta_0\otimes\nu_0\rangle = \langle \hat{\mathcal{L}}h, \hat{\mu}\rangle, \tag{25}$$

and

$$\langle 1, \hat{\kappa}_2\rangle \leq p\langle 1, \delta_0\otimes\nu_0\rangle, \tag{26}$$
$$\nu_0 \leq \mu_0 \tag{27}$$

Then the complement of the set

$$\operatorname{supp}(\mu_0 - \nu_0^*)$$

is a weakly p-safe set and contains the largest strongly p-safe set with respect to $\preceq$.

The inequality $\nu_0 \leq \mu_0$ in (27) is to be understood as the following predicate: There is $\alpha \in \mathcal{M}_+(\mathcal{Y})$ such that $\nu_0 + \alpha = \mu_0$.

Proof: At the outset, we define

$$\bar{\mu} \equiv \frac{\hat{\mu}^*}{\nu_0^*(\mathcal{Y})}$$
$$\bar{\nu}_0 \equiv \frac{\nu_0^*}{\nu_0^*(\mathcal{Y})}$$
$$\bar{\kappa}_i \equiv \frac{1}{\nu_0^*(\mathcal{Y})}\hat{\kappa}_i^*, \quad i \in \{1,2\},$$

and $\bar{\kappa} \equiv \bar{\kappa}_1 + \bar{\kappa}_2$. Subsequently, the triple $(\bar{\mu}, \bar{\nu}_0, \bar{\kappa})$ solves the evolution equation for $\hat{\mathcal{L}}$, and $\bar{\nu}_0$ is a probability measure. Hence by Corollary 1,

$$\begin{aligned}&\mathbb{P}^{\delta_0\otimes\bar{\nu}_0}[\tau_{\mathbb{R}_+\times U} < \zeta_{[0,T]\times S}] = \bar{\kappa}([0,T)\times U)\\ =&\bar{\kappa}_1([0,T)\times U) + \bar{\kappa}_2([0,T)\times U) \leq \bar{\kappa}_2([0,T]\times\mathcal{Y}) \leq p.\end{aligned} \tag{28}$$

Let $A \equiv (\operatorname{supp}(\mu_0 - \nu_0^*))^c$. It follows that $(\mu_0 - \nu_0^*)(A) = 0$.

[1]The optimisation is over the cone of positive measures.

Next, we show that A is weakly p-safe

$$\begin{aligned}\chi^{\mu_0}(A) &= \int_A \mathbb{P}^{(0,y)}[\tau_{\mathbb{R}_+\times U} < \zeta_{[0,T]\times S}]\mu_0(dy)\\ &= \int_A \mathbb{P}^{(0,y)}[\tau_{\mathbb{R}_+\times U} < \zeta_{[0,T]\times S}]\nu_0^*(dy)\\ &= \nu_0^*(\mathcal{Y})\int_A \mathbb{P}^{(0,y)}[\tau_{\mathbb{R}_+\times U} < \zeta_{[0,T]\times S}]\bar{\nu}_0(dy)\\ &\leq \mathbb{P}^{\delta_0\otimes\bar{\nu}_0}[\tau_{\mathbb{R}_+\times U} < \zeta_{[0,T]\times S}] \leq p.\end{aligned}$$

We have used the fact that $\nu_0^*(\mathcal{Y}) \leq \mu_0(\mathcal{Y}) = 1$ and (28).

Now, we will prove by contradiction that A contains the largest strongly p-safe set. Suppose that there is a strongly p-safe set $A' \subset S \setminus U$ such that $A \cap A' = \emptyset$ and $\mu_0(A') \neq 0$. Let $\alpha \equiv \mu_0|_{A'} - \nu_0^*|_{A'}$. Notice that $\alpha \in \mathcal{M}_+(\mathcal{Y})$, $\alpha(\mathcal{Y}) > 0$, and $\nu_0^* + \alpha \leq \mu_0$. By Proposition 1,

$$\int_{\mathcal{Y}} \mathbb{P}^{(0,y)}[\tau_{\mathbb{R}_+\times U} < \zeta_{[0,T]\times S}]\alpha(dy) \leq p\int_{\mathcal{Y}} \alpha(dy).$$

For the initial probability measure α, we construct the occupation measure $\hat{\mu}'$ and the hitting measure $\hat{\kappa}'$ as in (14) and (15). We define $\hat{\kappa}'_i$, $i \in \{1,2\}$ as in (22). We observe that $\hat{\mu}', \hat{\kappa}'_1, \hat{\kappa}'_2$ satisfy (23) and $(\hat{\mu}', \alpha, \hat{\kappa}'_1 + \hat{\kappa}'_2)$ solves the evolution equation for $\hat{\mathcal{L}}$. Since all the constraints in the optimisation problem are linear, $(\hat{\mu}, \nu_0, \hat{\kappa}_1, \hat{\kappa}_2) = (\hat{\mu}^* + \hat{\mu}', \nu^* + \alpha, \hat{\kappa}_1^* + \hat{\kappa}'_1, \hat{\kappa}_2^* + \hat{\kappa}'_2)$ solves (23) to (27). But this is a contradiction as $(\alpha+\nu_0^*)(\mathcal{Y}) > \nu_0^*(\mathcal{Y})$. ∎

Corollary 2: Suppose that the premises of Theorem 2 hold with the constraint $\int d\hat{\kappa}_2 \leq p\int d\hat{\nu}_0$ substituted by $\int d\hat{\kappa}_2 > p\int d\hat{\nu}_0$. Then the complement of the set

$$V = \mathrm{supp}(\mu_0 - \nu_0^*)$$

is a weakly p-unsafe set and contains the largest strongly p-unsafe set.

The virtue of Corollary 2 is that all p-unsafe states are excluded from the found set V. In other words, contrary to Theorem 2, it under-approximates the p-safe states, which is a desired property of any safety-verification algorithm.

Let R be the weakly p-safe set in Theorem 2 and R' be the weakly p-unsafe set in Corollary 2, then the following inclusions show that $R \setminus R'$ is a subset of the largest strongly p-safe set A

$$R \setminus R' \subset R \setminus A^c \subset S \setminus A^c = A.$$

In conclusion, we have the following evaluation of the largest strongly p-safe set

$$R \setminus R' \subset A \subset R.$$

The main result below shows that p-safety is equivalent to the existence of the quadruple $(\hat{\mu}^*, \nu_0^*, \hat{\kappa}_1^*, \hat{\kappa}_2^*) \in \mathcal{O}$ of measures solving the optimisation problem (24).

Proposition 5: There exists a measure ν_0^* solving (24) to (27). Furthermore, there is a unique density ρ^* such that

$$d\nu_0^* = \rho^* d\mu_0. \tag{29}$$

Proof: First, we observe that the optimisation problem is feasible. To this end, we let $\nu_0 = \mu_0|_A$, and define $\hat{\mu}$ as in (14) with μ_0 substituted by ν_0, $\hat{\kappa}_1$, and $\hat{\kappa}_2$ as in (22) with $\hat{\kappa}$ defined in (15).

We will show that there exists a measure ν_0^* solving (24) to (27). Since $\nu_0 \leq \mu_0$ (there is $\alpha \in \mathcal{M}_+(\mathcal{Y})$ such that $\nu_0 + \alpha = \mu_0$), ν_0 is absolutely continuous with respect to μ_0. Therefore, by Radon-Nikodym theorem, there is a unique density ρ, a nonnegative measurable function on $\mathcal{Y}$ such that $d\nu_0 = \rho d\mu_0$. Specifically $\rho \leq 1$ as $\nu_0 \leq \mu_0$. Consider a sequence of measures (ν_0^j) satisfying $\lim_{j\to\infty} \nu_0^j(\mathcal{Y}) = v^*$. To each measure ν_0^j, we associate the density ρ^j. By dominated convergence theorem, there is a density ρ^* such that

$$\lim_{j\to\infty} \nu_0^j(\mathcal{Y}) = \lim_{j\to\infty} \int \rho^j d\mu_0 = \int \rho^* d\mu_0.$$

We define ν_0^* by $d\nu_0^* = \rho^* d\mu_0$. ∎

Theorem 3: Let μ_0 be an initial measure on $\mathcal{Y}$ with $\mathrm{supp}(\mu_0) \subseteq \mathcal{Y}$. A set A is a maximal p-safe set if and only if there exists a quadruple $(\hat{\mu}^*, \nu_0^*, \hat{\kappa}_1^*, \hat{\kappa}_2^*) \in \mathcal{O}$ that solves the optimisation problem given by (24) to (27), and $A \subset (\mathrm{supp}(\mu_0 - \nu_0^*))^c$.

Proof: The sufficiency follows from Theorem 2, and the necessity from Proposition 5. ∎

As the consequence of Proposition 5, we use the density ρ^* to characterise the largest strongly p-safe set. Specifically for an $A \subset S$, if $\rho^*|_A = 1$ then A is weakly p-safe.

Later in the paper, we use the moment method of [20] to compute an approximation of ν_0^*. Suppose for a while that $\mathcal{Y} \subset \mathbb{R}^n$. Consequently, the sequences of reals are computed corresponding to the moments of the finite Borel measures $\hat{\mu}^*, \nu_0^*, \hat{\kappa}_1^*$, and $\hat{\kappa}_2^*$. Let the sequence $(y_\alpha^{\nu_0})_{\alpha\in N^n}$ be the moments of ν_0^*, and $(y_\alpha^{\mu_0})_{\alpha\in N^n}$ be the moments of μ_0. If both sequences of moments are finite, then from (29), ρ^* is a polynomial function $\rho^*(X) = \sum_{\gamma\in\mathbb{N}^n} \rho_\gamma^* X^\gamma$ of degree d, where the coefficients ρ_γ^* are uniquely determined by the system of linear equations

$$y_\alpha^{\nu_0} = \sum_{\gamma+\beta=\alpha} \rho_\gamma^* y_\beta^{\mu_0}.$$

The results derived in this section can be applied for the processes with (extended) generators satisfying Assumption 1. The list includes diffusion and jump diffusion processes. Nonetheless, Assumption 1 does not hold for processes with forced jumps. In the next sections, we will show how to circumvent this problem for piecewise-deterministic Markov processes.

VII. Piecewise-deterministic Markov Processes

A piecewise-deterministic Markov process, in short PDMP, is a stochastic hybrid process, whose dynamics represent alternation of continuous dynamics captured by deterministic differential equations with spontaneous and forced stochastic jumps [12]. The jumps are characterised by stochastic time of jumps and stochastic jump-destinations.

We consider a countable set Q of discrete states, a function $d : Q \to \mathbb{N}$, and $\mathcal{X} : Q \to \bigcup_{j\in Q} \{j\}\times\mathcal{P}(\mathbb{R}^{d(j)})$ a multivalued map assigning to each discrete state $j \in Q$ an open subset of $\{j\} \times \mathbb{R}^{d(j)}$. A hybrid state space of a PDMP is the set

$$\mathcal{Y} \equiv \bigcup_{j\in Q} \mathcal{X}(j) = \{(j,x) \in \mathcal{X}(j) |\ j \in Q\}.$$

A vector field f on the hybrid state space $\mathcal{Y}$ is a map

$$f : \mathcal{Y} \to \bigcup_{j\in Q} \{j\} \times \mathbb{R}^{d(j)}$$

with the section property $\pi_1 f = \mathrm{id}$, where $\pi_1(j,x) = j$, and id is the identity map on Q. In other words, f assigns to each hybrid state $y = (j,x)$ a vector $f(y) \in \{j\} \times \mathbb{R}^{d(j)}$. For an open interval $(-\epsilon, \epsilon) \subset \mathbb{R}$ the flow of f is a map

$$\phi : (-\epsilon, \epsilon) \times \mathcal{Y} \to \mathcal{Y}$$

such that $\phi(0,y) = y$

$$\frac{d}{dt}\phi(t,y) = f(\phi(t,y)), \quad \text{for } t \in (-\epsilon,\epsilon). \tag{30}$$

Let Γ be the set of points on the boundary $\partial\mathcal{Y}$ of $\mathcal{Y}$, which are reached by the flow of the vector field f

$$\Gamma \equiv \{y \in \partial\mathcal{Y} |\ \exists (t,y') \in \mathbb{R}_+ \times \mathcal{Y}, y = \phi(t,y')\}.$$

Let

$$\overline{\mathcal{Y}} \equiv \mathcal{Y} \cup \Gamma.$$

We are now ready to define a piecewise-deterministic Markov process.

Definition 6 (PDMP): A piecewise-deterministic Markov process is a collection

$$((Q,d,\mathcal{X}), f, \nu_0, \lambda, K),$$

where

- $(Q,d,\mathcal{X})$, and f are as described above;
- $\nu_0 : \mathcal{B}(\mathcal{Y}) \to [0,1]$ is an initial probability measure;
- $\lambda : \mathcal{Y} \to \mathbb{R}_+$ is a jump rate that defines the survivor function F of the ith jump after the jump at time T_{i-1} from the hybrid state y_{i-1} as follows

$$F(y_{i-1},t) = \begin{cases} \exp\left(-\int_0^t \lambda(\phi(y_{i-1},\tau))d\tau\right) & \text{if } t < t^*(y_{i-1}) \\ 0 & \text{if } t \geq t^*(y_{i-1}), \end{cases}$$

  where $t^* : \mathcal{Y} \to \mathbb{R}_+ \cup \{\infty\}$ is the exit time defined by

$$t^*(y) = \inf\{t > 0 |\ \phi(y,t) \notin \mathcal{Y}\};$$

- $K : \mathcal{B}(\mathcal{Y}) \times \overline{\mathcal{Y}} \to [0,1]$ is a transition measure. At any point $y \in \mathcal{Y}$, $K(A,y)$ is the probability that the jump is onto the set A.

The execution of a PDMP consists of a deterministic process (X_t) corresponding to the solution of ODE starting at y_{i-1}, alternating with stochastic jumps at time T_i defined by the survivor function $F(y_{i-1},t)$. The post-jump location y_i is described by the transition measure $K(dy_i, y_{i-1})$.

The extended generator of a PDMP is characterised by the following theorem.

Theorem 4 (Theorem 11.2.2 in [12]):
Let $((Q,d,\mathcal{X}),\xi,\nu_0,\lambda,K)$ be a PDMP. The restriction $g|_{\mathcal{Y}}$ of a measurable function $g : \overline{\mathcal{Y}} \to \mathbb{R}$ to $\mathcal{Y}$ belongs to the domain $D(\mathcal{L})$ of the extended generator $\mathcal{L}$ if

- For each $y \in \mathcal{Y}$, the function $t \mapsto g(\phi(t,y))$ is absolutely continuous on the open interval $(0, t^*(y))$.
- For each y in the active boundary Γ,

$$g(y) = \int_{\mathcal{Y}} g(z) K(dz; y). \tag{31}$$

- For each $y \in \mathcal{Y}$, and $t \geq 0$,

$$E^y\left[\sum_{i\in N_t} |g(X_{T_i}) - g(X_{T_{i^-}})|\right] < \infty, \tag{32}$$

  where $X^y_{T_{i^-}}$ is the limit $\lim_{t\uparrow T_i} X^y_t$ from the left.

Furthermore, for $g \in \mathcal{D}(\mathcal{L})$, the extended generator $\mathcal{L}$ is defined by

$$\mathcal{L}g(y) \equiv \langle f(y), \nabla g(y)\rangle + \lambda(y)\int_{\mathcal{Y}} (g(z) - g(y)) K(dz; y). \tag{33}$$

Example 2: Consider a PDMP with two modes $K = \mathbb{Z}_2$. Firstly, denote the open box $]-1,1[\times]-1,1[\subset \mathbb{R}^2$ by $\overset{\circ}{B}$, and its closure in $\mathbb{R}^2$ by B. Suppose that $d : \mathbb{Z}_2 \to \{2\}$, and $\mathcal{X} : \mathbb{Z}_2 \to \{\overset{\circ}{B}\}$. As a consequence, the hybrid state space is $\mathcal{Y} = \mathbb{Z}_2 \times \overset{\circ}{B}$.

Let the vector field on $\mathcal{Y}$ be denoted by

$$f_0(x) \equiv f(0,x) \quad \text{and} \quad f_1(x) \equiv f(1,x).$$

For explicit forms of functions f_0 and f_1 see (40) in Subsection IX.

Suppose that the jump rate is $\lambda(v,x) = \lambda_v$ (λ_0 and λ_1 are constant on the continuous states). Consider deterministic jumps from the mode v to $v+1 \mod 2$ ($0 \mapsto 1$, $1 \mapsto 0$) defined by maps $j_v : B \to \overset{\circ}{B}$. Subsequently, the transition measure is

$$\begin{aligned} &K(dy'; y) = Q(dv' \times dx'; (v,x)) \\ &= \left(\delta_{j_0(x)}(dx') I_{\{v=0\}} + \delta_{j_1(x)}(dx') I_{\{v=1\}}\right) \delta_{(v+1 \bmod 2)}(dv'). \end{aligned}$$

Finally, suppose that the initial measure is concentrated at a point $(v_0, \bar{x}_0) \in \mathcal{Y}$, $\nu_0 = \delta_{(v_0,\bar{x}_0)}$.

The realisations of the PDMP $((Q,d,\mathcal{X}), f, \nu_0, \lambda, K)$ for the initial state $(v_0, \bar{x}_0)$ are of the form

$$x_t(\omega) = (v, \phi_v(t, \bar{x}_i)) \text{ for } T_{i-1}(\omega) \leq t < T_i(\omega),$$

where ϕ_v is the flow map of the vector field f_v. At each time T_i, there is a jump from $(v, \phi_v(T_i(\omega), x_i))$ to

$$(v+1 \bmod 2, h_v(\phi_v(T_i(\omega), x_i))).$$

The time of the i^{th} jump is

$$\begin{aligned} T_0(\omega) &= 0 \\ T_i(\omega) &= T_{i-1}(\omega) + S_i(\omega),\ i \in \mathbb{N}, \end{aligned}$$

where S_i are independent random variables of exponential distribution with the intensity $\lambda_{(i \bmod 2)}$.

Using the notation $g_v(x) \equiv g(v,x)$, the extended generator in (33) is expressed as

$$\begin{aligned} \mathcal{L}g(v,x) &= \frac{\partial g_0}{\partial x}(x) f_0 \delta_0(v) + \frac{\partial g_1}{\partial x}(x) f_1 \delta_1(v) \\ &+ \lambda_v \left(g_1 \circ j_0(x) I_{\{v=0\}} + g_0 \circ j_1(x) I_{\{v=1\}} - g_v(x)\right). \end{aligned}$$

Finally, (31) takes the form: For any $x \in \Gamma \subset \partial B$ and any $t \in \mathbb{R}_+$

$$f_0(t,x) = f_1(t,h_0(x))$$
$$f_1(t,x) = f_0(t,h_1(x)),$$

where Γ in plain words is the subset of the boundary of $\mathcal{Y}$ that is reached by the realisations.

VIII. Occupation Measures for PDMPs

We observe that the domain of the extended generator of PDMP does not satisfy Assumption 1. Specifically, relation (31) prohibits $\mathcal{D}(\mathcal{L})$ from being closed under multiplication. Our strategy is to extend the domain of the operator $\mathcal{L}$ such that Assumption 1 is satisfied.

We define an operator $\mathcal{C}h(t,y) \equiv \int_{\mathcal{Y}} h(t,z)K(dz;y) - h(t,y),\ y \in \Gamma$, and an increasing predictable process

$$p_t = \sum_{T_i \le t} I_{[x_{T_i^-} \in \Gamma]}.$$

Using the PDMP differential formula [12, Sec. 31], we conclude that the process

$$M^h_{t\wedge\tau} \equiv h(t\wedge\tau, X_t) - h(0,X_0) - \int_0^{t\wedge\tau} \hat{\mathcal{L}}h(s,X_s)ds - \int_0^{t\wedge\tau} \mathcal{C}h(s,X_{s^-})dp(s),$$

where $\hat{\mathcal{L}} = \mathcal{L} + \frac{\partial}{\partial t}$ and $\mathcal{L}$ is given by (33), is a martingale for h absolutely continuous. As a consequence,

$$\mathbb{E}^{(0,y)}[h(t\wedge\tau, X_{t\wedge\tau})] = \mathbb{E}^{(0,y)}[h(0,X_0)] + \mathbb{E}^{(0,y)}\int_0^{t\wedge\tau} \hat{\mathcal{L}}h(s,X_s)ds + \mathbb{E}^{(0,y)}\int_0^{t\wedge\tau} \mathcal{C}h(s,X_{s^-})dp(s).$$

We have already introduced the hitting and the occupation measures in Subsection V-B, we will in addition define a measure $\hat{\rho}$ by

$$\hat{\rho}(B\times C) \equiv \int_{\mathbb{R}_+\times\mathcal{Y}} \rho((b,y), B\times C)\hat{\mu}_0(dy).$$

where

$$\rho((b,y), B\times C)) \equiv \mathbb{E}^{(b,y)} \int_{[0,t\wedge\tau]} I_{B\times C}(s,X_{s^-})dp(s) = \sum_{T_i\le\tau} I_{B\times(C\cap\Gamma)}(T_i, X_{T_i^-}).$$

Intuitively, $\hat{\rho}$ is the part of the occupation measure that corresponds to the jumps.

Equipped with the quadruple of measures $(\hat{\mu}, \mu_0, \hat{\kappa}, \hat{\rho})$ with

$$\mathrm{supp}(\mu_0) \subset S\setminus U, \tag{34a}$$
$$\mathrm{supp}(\hat{\kappa}) \subset \{T\}\times \mathrm{cl}(S\setminus U) \cup [0,T]\times S^c \cup [0,T]\times U, \tag{34b}$$
$$\mathrm{supp}(\hat{\mu}) \subset [0,T]\times \mathrm{cl}(S\setminus U), \tag{34c}$$
$$\mathrm{supp}(\hat{\rho}) \subset [0,T]\times(\mathrm{cl}(S\setminus U)\cap\Gamma) \tag{34d}$$

the PDMP is characterised by the following relation

$$\langle h, \hat{\kappa}\rangle = \langle h, \hat{\mu}_0\rangle + \langle \hat{\mathcal{L}}h, \hat{\mu}\rangle + \langle \mathcal{C}h, \hat{\rho}\rangle. \tag{35}$$

In consequence, we are able to modify the results in Section VI to encounter for PDMPs. Specifically, Theorem 2 and Corollary 2 are reformulated as follows.

Theorem 5: Let μ_0 be an initial (probability) measure on $\mathcal{Y}$ with $\mathrm{supp}(\mu_0) \subseteq S$. Suppose that the quadruple $(\hat{\mu}^*, \delta_0\otimes\nu_0^*, \hat{\kappa}_1^*, \hat{\kappa}_2^*, \hat{\rho}^*)$ of measures on $[0,T]\times\mathcal{Y}$ is a solution to the conic optimisation

$$v^* = \sup\{\nu_0(\mathcal{Y})|\ \hat{\mu}, \delta_0\otimes\nu_0, \hat{\kappa}_1, \hat{\kappa}_2, \hat{\rho} \in \mathcal{M}_+([0,T]\times\mathcal{Y})\}$$

such that $\forall h \in A_b([0,T]\times\mathcal{Y})$

$$\langle h, \hat{\kappa}_1\rangle + \langle h, \hat{\kappa}_2\rangle - \langle h, \delta_0\otimes\nu_0\rangle = \langle \hat{\mathcal{L}}h, \hat{\mu}\rangle + \langle \mathcal{C}h, \hat{\rho}\rangle, \tag{36}$$

and (23), (26), (27) together with (34d) . Then the complement of the set

$$\mathrm{supp}(\mu_0 - \nu_0^*)$$

is a weakly p-safe set and contains the largest strongly p-safe set with respect to $\preceq$.

Corollary 3: Suppose that the premises of Theorem 5 hold with the constraint $\int d\hat{\kappa}_2 \le p\int d\hat{\nu}_0$ substituted by $\int d\hat{\kappa}_2 > p\int d\hat{\nu}_0$. Then the complement of the set

$$\mathrm{supp}(\mu_0 - \nu_0^*)$$

is a weakly p-unsafe set and contains the largest strongly p-unsafe set.

Example 3: We continue with Example 2. We use the extended generator of the time-space process

$$\hat{\mathcal{L}}h(t,(v,x)) = \mathcal{L}h_t(v,x) + \frac{\partial h(t,(v,x))}{\partial t}. \tag{37}$$

In (37), we have used the notation $h_t(v,x) \equiv h(t,(v,x))$ to indicate that the variable t is seen as a parameter.

Consequently, noticing that the occupation measure takes the form

$$\hat{\mu}(dt\times dy) = \hat{\mu}(dt\times dv\times dx) = \mu^0(dt\times dx)\delta_0(dv) + \mu^1(dt\times dx)\delta_1(dv),$$

the term in (35) that involves the generator in the optimisation formulation takes the form

$$\langle \hat{\mathcal{L}}h, \hat{\mu}\rangle = \int_{[0,T]\times\mathcal{Y}} \hat{\mathcal{L}}h(t,(v,x))\hat{\mu}(dt\times dv\times dx) \tag{38}$$

$$= \int_{[0,T]\times\overset{\circ}{B}} \left(\frac{\partial h_0}{\partial t} + \frac{\partial h_0}{\partial x}f_0 + \lambda_0(h_1\circ\mathrm{id}\times j_0 - h_0)\right)d\mu^0 + \int_{[0,T]\times\overset{\circ}{B}} \left(\frac{\partial h_1}{\partial t} + \frac{\partial h_1}{\partial x}f_1 + \lambda_1(h_0\circ\mathrm{id}\times j_1 - h_1)\right)d\mu^1. \tag{39}$$

In (38), we have suppressed the arguments (t,x) and denoted $h_v(t,h_w(x))$ by $(h_v\circ\mathrm{id}\times j_w)(t,x)$, $w \in \mathbb{Z}_2$. Again noticing that

$$\hat{\rho}(dt\times dv\times dx) = \rho^0(dt\times dx)\delta_0(dv) + \rho^1(dt\times dx)\delta_1(dv),$$

the last term in (35) becomes

$$\langle \mathcal{C}h, \hat{\rho} \rangle = \int_{[0,T]\times\mathcal{Y}} \mathcal{C}h(t,(v,x))\hat{\rho}(dt \times dv \times dx)$$
$$= \sum_{v\in\mathbb{Z}_2} (h(t,(v+1, j_v(x))) - h(t,(v,x)))\rho^v(dt \times dx).$$

Notice that if $\lambda_0 = \lambda_1 = 0$, $f_1 = 0$, and $p = 1$. The p-safety verification of this PDMP corresponds to the verification of the dynamical system $\dot{x} = f_0(x)$; whereas, if the continuous state space is $\mathbb{R}$, $\lambda_0 = \lambda_1$, $f_0 = f_1 = 0$, and $j_0(x) = j_1(x) = x+1$, it corresponds to p-safety verification of a Poisson process with intensity λ_0.

IX. Example

This section provides a numerical validation of the proposed method for finding p-safe initial states of a PDMP with three different jump rates, similar to the PDMP in [28]. Results of Monte Carlo simulations are provided to give an approximation of the largest strong p-safe set. The simulation-based results are subsequently compared to the weakly p-safe set found with our optimization-based solution. The optimization-based solution provides an inner approximation of the weakly p-safe set to ensure soundness of the algorithm; however, the Monte Carlo simulations provide an approximation of the strongly p-safe set. Therefore, the approximate of the weakly p-safe set includes the strongly p-safe set identified by Monte Carlo simulations.

A. Monte Carlo Simulations

Similar to Example 2, we consider a PDMP with two modes $K = \mathbb{Z}_2$ and vector field $f(v,x) = (v, f_v)$ given by

$$f_0(x) = \begin{bmatrix} 2x_1 - x_2^2 + 2 \\ x_2 \end{bmatrix} \quad \text{and} \quad f_1(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{40}$$

where $x = (x_1, x_2)$, and both jump maps are chosen to be identity.

The state space of the system is $S = \{0,1\} \times B$, where $B = [-1,1] \times [-1,1] \subset \mathbb{R}^2$. The jump rates are chosen to be constant on the continuous states

$$\lambda(0,x) = \lambda(1,x) = \lambda.$$

Three different jump rates ($\lambda = 0$, $\lambda = 5$, and $\lambda = 100$) are considered, where $\lambda = 0$ corresponds to no jumps, $\lambda = 100$ corresponds to very frequent jumps, and $\lambda = 5$ is an intermediate value. This is seen from a realization of the time between jumps, which can be found numerically from

$$\Delta t_{\text{jump}} = -\frac{\log(X)}{\lambda},$$

where X is a uniformly distributed random variable on the interval $[0,1]$.

In the numerical study, the final time is set to $T = 10$, and the set of unsafe states is defined as

$$U = \{0,1\} \times \{x \in B |\ x_1 + x_2 - 1.7 > 0\}.$$

1) Jump rate $\lambda = 0$: When the jump rate of a PDMP is zero, no jumps happen between the discrete states of the system, i.e., the system is deterministic. As a consequence, an initial state is either safe or unsafe with probability one. A simulation of the two vector fields f_0 and f_1 without any jumps is shown in Fig. 2. The green diagonal hatched area illustrates the largest strongly p-safe set for $\lambda = 0$.

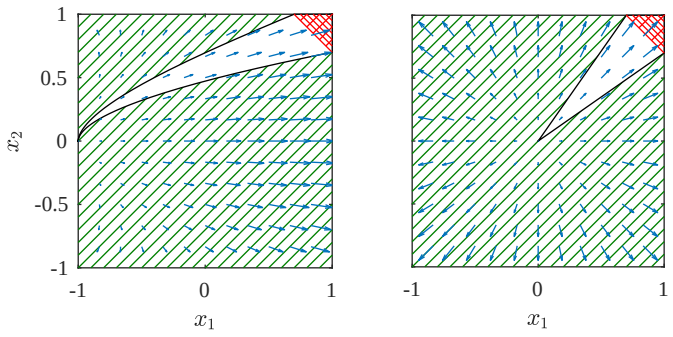

Fig. 2. Phase plot (blue arrows) of the two vector fields f_0 (left) and f_1 (right). The red area (upper right corner) is the set of unsafe states U, the green diagonal hatched area is the set of safe initial states, when no jumps occur.

2) Jump rate $\lambda = 5$: When the jump rate of the PDMP is larger than zero, then jumps occur between the discrete states of the system. In this case, an approach similar to [29] is used for the simulation of PDMPs. We use a discretisation of the state space into a uniform grid of cells of size $n_x \times n_y$ ($n_x, n_y \in \mathbb{N}$), with each cell defined as

$$B_{i,j} = \Big\{ (x,y) \in B |\ \frac{i-1}{n_x}(\overline{x} - \underline{x}) \le x - \underline{x} \le \frac{i}{n_x}(\overline{x} - \underline{x}), \quad \frac{j-1}{n_y}(\overline{y} - \underline{y}) \le y - \underline{y} \le \frac{j}{n_y}(\overline{y} - \underline{y}) \Big\}$$

for $i = 1, \ldots, n_x$ and $j = 1, \ldots, n_y$ where $\underline{x} = \min_{(x,y)\in B} x$ and $\overline{x} = \max_{(x,y)\in B} x$ (and similarly for $\underline{y}$, $\overline{y}$).

The Monte Carlo simulation is performed according to Algorithm 1. For a set X, `Rand(X)` returns a point $x \in X$ given by a uniform distribution on X and `sim`$(f_k, [0, t_{\text{jump}}], x_0)$ returns the trajectory $x(t)$ of the differential equation

$$\dot{x} = f_k(x)$$

for $t \in [0, t_{\text{jump}}]$ with $x(0) = x_0$.

The Monte Carlo simulation is performed by partitioning the state space uniformly into 10,000 cells, i.e., $n_x = n_y = 100$ and simulating 100 trajectories from each cell ($N = 100$), with initial conditions given by a uniform distribution on each cell. The result of the Monte Carlo simulation is shown in Fig. 3 as a contour plot, where the contour lines indicate probabilities of reaching the unsafe set U. The contour lines indicate the probability of reaching the unsafe set, which is approximated by the fraction of simulations that reached the unsafe set in the 100 simulations. The contour line closest to the upper right corner indicates a probability of 0.9 for reaching U, and this probability decreases by 0.1 per contour line.

Algorithm 1 Monte Carlo Simulation

procedure
  for $k_0 \in K$ **do**
    for $(i,j) \in \{(i,j) \in \mathbb{N}^2 |\ i \le n_x, j \le n_y\}$ **do**
      for $l = 1, \ldots, N$ **do**
        $t_0 \leftarrow 0$
        $k \leftarrow k_0$
        $x_0 \leftarrow \texttt{rand}(B_{i,j})$
        stopFlag $\leftarrow$ `False`
        while stopFlag==`False` **do**
          $t_{\text{jump}} = -\frac{\log(\texttt{rand}([0,1]))}{\lambda}$
          $x(t) = \texttt{sim}(f_k, [0, t_{\text{jump}}], x_0)$
          $t_U = \min_t x(t) \in U$
          $t_S = \max_t x([0,t]) \subset B$
          if $t_U \le \min(t_S, T - t_0)$ **then**
            stopFlag $\leftarrow$ `True`
            Unsafe
          else if $t_S < \min(t_{\text{jump}}, T - t_0)$ **then**
            stopFlag $\leftarrow$ `True`
            Safe
          else if $t_0 + t_{\text{jump}} > T$ **then**
            stopFlag $\leftarrow$ `True`
            Safe
          else
            $t_0 \leftarrow t_0 + t_{\text{jump}}$
            $k \leftarrow \texttt{mod}(k+1, 2)$
            $x_0 \leftarrow x(t_{\text{jump}})$

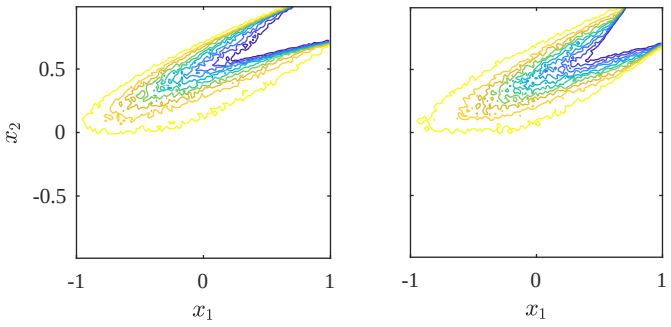

Fig. 3. Contour plots showing the probability of reaching the unsafe set U from initial discrete state $k = 1$ (left) and initial discrete state $k = 2$ (right), when the jump rate is $\lambda = 5$. The simulations are performed with $n_x = n_y = 100$ and $N = 100$, i.e., 1,000,000 simulations are performed in total with a combined simulation time of 2 hours.

3) Jump rate $\lambda = 100$: Monte Carlo simulations are also performed for the PDMP with jump rate $\lambda = 100$ following Algorithm 1, according to the same approach as for the case where $\lambda = 5$. The result of the Monte Carlo simulation is shown in Fig. 4 in terms of a contour plot.

It is seen in Fig. 4 that the initial discrete state of the system has little impact on the safety of the initial state of the system when the jump rate is very high ($\lambda = 100$). It should be observed that the simulation time is very short, at the time between jumps is very short.

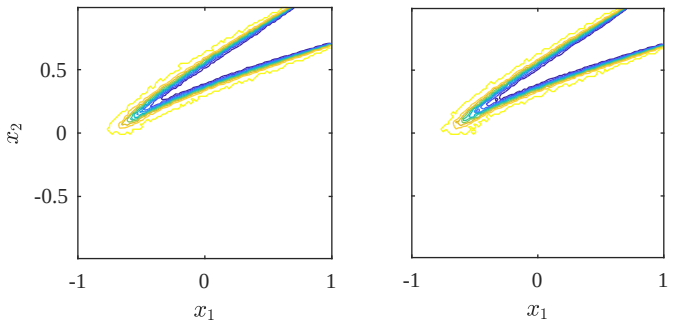

Fig. 4. Contour plots showing the probability of reaching the unsafe set U from initial discrete state $k = 1$ (left) and initial discrete state $k = 2$ (right), when the jump rate is $\lambda = 100$. The simulations are performed with $n_x = n_y = 100$ and $N = 100$, i.e., 1,000,000 simulations are performed in total with a total simulation time of 19 hours.

B. Optimization-Based Safety Analysis

The following explains how a weakly p-safe set is found by use of the numerical optimization problem presented in Theorem 5, with $p = 0.1$.

The optimization problem is modified, to be implemented as a numerical optimization problem that can be solved in MATLAB using GloptiPoly 3. To this end, we impose additional assumptions on the vector field f in the PDMP formulation to be real polynomial, and S and U to be semi-algebraic (sets generated by polynomials). We also suppose that the set S is bounded.

Since the measure $\hat{\kappa}_1$ has support on the union of two subsets, it is split into two measures $\hat{\kappa}_{1,1}$ and $\hat{\kappa}_{1,2}$ with supports on each of the subsets. In addition, the initial measure μ_0 is chosen to be the Lebesgue measure on $\mathcal{Y}$. Thus, we attempt to solve the optimization problem

$$\sup \nu_0(\mathcal{Y})$$

such that

$$\langle h, \hat{\kappa}_{1,1} + \hat{\kappa}_{1,2} \rangle + \langle h, \hat{\kappa}_2 \rangle - \langle h, \delta_0 \otimes \nu_0 \rangle = \langle \hat{\mathcal{L}} h, \hat{\mu} \rangle + \langle \mathcal{C} h, \hat{\rho} \rangle$$

and

$$\begin{aligned}
&\langle 1, \hat{\kappa}_2 \rangle \le p \langle 1, \delta_0 \otimes \nu_0 \rangle, \\
&\nu_0 + \nu_0^c = \mu_0, \\
&\text{supp}(\hat{\mu}) \subset [0,T] \times \text{cl}(S \setminus U), \\
&\text{supp}(\nu_0) \subset S \setminus U, \\
&\text{supp}(\nu_0^c) \subset S \setminus U, \\
&\text{supp}(\hat{\kappa}_{1,1}) \subset \{T\} \times \text{cl}(S \setminus U), \\
&\text{supp}(\hat{\kappa}_{1,2}) \subset [0,T] \times S^c, \\
&\text{supp}(\hat{\kappa}_2) \subset [0,T] \times U.
\end{aligned}$$

where $\hat{\mu}, \delta_0 \otimes \nu_0, \delta_0 \otimes \nu_0^c, \hat{\kappa}_{1,1}, \hat{\kappa}_{1,2}, \hat{\kappa}_2 \in \mathcal{M}_+([0,T] \times \mathcal{Y})$.

The above optimization problem should hold for all $h \in A_b([0,T] \times \mathcal{Y})$; however, in the optimization, the problem is relaxed - truncated such that measures with moments up to degree 10 are used in the implementation.

The optimization problem is formulated in MATLAB using GloptiPoly 3 and YALMIP, and the resulting SDP is solved using SeDuMi. The standard absolute accuracy for checking

feasibility of constraints (0.001) was used when solving the SDP problem.

The sets obtained from the optimization are shown in Fig. 5, Fig. 6, and Fig. 7. It takes between 80 s and 130 s to solve each of the three optimization problems on a standard laptop PC.

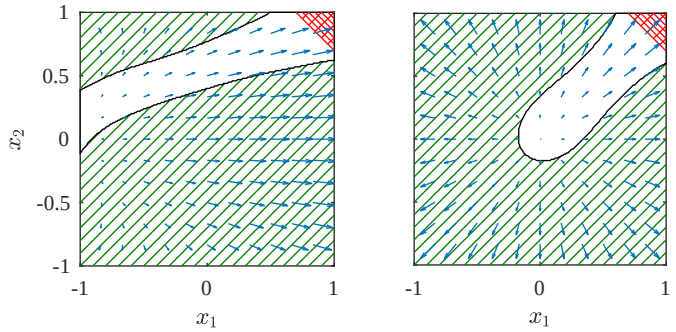

Fig. 5. Phase plot (blue arrows) of the two vector fields. The red diagonal cross hatched area is the set of unsafe states U, the green diagonal hatched area is the weakly 0.1-safe set with $\lambda = 0$.

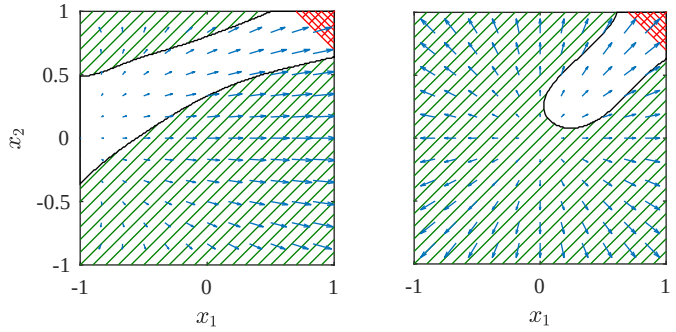

Fig. 6. Phase plot (blue arrows) of the two vector fields. The red diagonal cross hatched area is the set of unsafe states U, the green diagonal hatched area is the weakly 0.1-safe set with $\lambda = 5$.

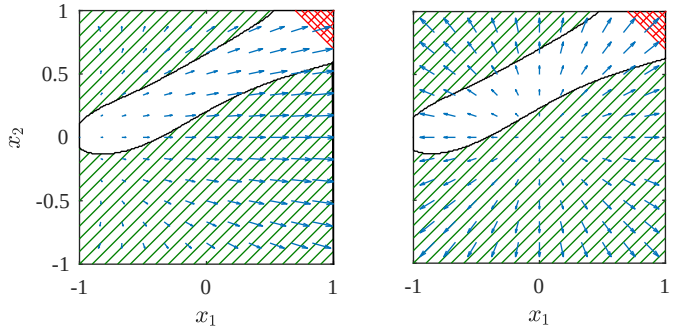

Fig. 7. Phase plot (blue arrows) of the two vector fields. The red diagonal cross hatched area is the set of unsafe states U, the green diagonal hatched area is the weakly 0.1-safe set with $\lambda = 100$.

X. Conclusions

We defined the concept of probabilistic safety (weak and strong p-safety). Our analysis started with the martingale problem associated to the underlying stochastic process. Using this, we derived the adjoint equation that relates the occupation measures and the hitting probabilities. Equipped with the adjoined equation, we formulated a linear infinite dimensional optimisation problem, which can be solved by the generalised moment method. We applied our method to probabilistic safety analysis of a piecewise-deterministic Markov process that can be thought of as a realisation of some stochastic hybrid system. A numerical example illustrating the method was provided.

References

[1] S. Prajna, "Barrier certificates for nonlinear model validation," *Automatica*, vol. 42, no. 1, pp. 117–126, 2006.

[2] R. Wisniewski, M. Svenstrup, A. S. Pedersen, and C. S. Steiniche, "Certificate for safe emergency shutdown of wind turbines," in *Proceedings of American Control Conference*, 2013, pp. 3667–3672.

[3] P. Wieland and F. Allgöwer, "Constructive safety using control barrier functions," in *Proceedings of the 7th IFAC Symposium on Nonlinear Control Systems*, 2007, pp. 462–467.

[4] M. Z. Romdlony and B. Jayawardhana, "Stabilization with guaranteed safety using control Lyapunov-barrier function," *Automatica*, vol. 66, no. C, pp. 39–47, Apr. 2016.

[5] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, Mar. 2001.

[6] D. Henrion, J. B. Lasserre, and J. Loefberg, "GloptiPoly 3: moments, optimization and semidefinite programming," *Optimization Methods and Software*, vol. 24, no. 4-5, pp. 761–779, 2009.

[7] M. Korda, D. Henrion, and C. N. Jones, "Inner approximations of the region of attraction for polynomial dynamical systems," in *Proceedings of the 9th IFAC Symposium on Nonlinear Control Systems*, 2013, pp. 534–539.

[8] D. Henrion and M. Korda, "Convex computation of the region of attraction of polynomial control systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 297–312, February 2014.

[9] T. G. Kurtz and R. H. Stockbridge, "Existence of Markov controls and characterization of optimal Markov controls," *SIAM Journal on Control and Optimization*, vol. 36, no. 2, pp. 609–653, 1998.

[10] A. G. Bhatt and R. L. Karandikar, "Invariant measures and evolution equations for Markov processes characterized via martingale problems," *The Annals of Probability*, vol. 21, no. 4, pp. 2246–2268, 1993.

[11] M. J. Cho and R. H. Stockbridge, "Linear programming formulation for optimal stopping problems," *SIAM Journal on Control and Optimization*, vol. 40, no. 6, pp. 1965–1982, 2002.

[12] M. H. A. Davis, *Markov models and optimization*, ser. Monographs on Statistics and Applied Probability. Chapman & Hall, London, 1993, vol. 49.

[13] J.-P. Aubin, *Viability Theory*, ser. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1991.

[14] S. Prajna, A. Jadbabaie, and G. J. Pappas, "A framework for worst-case and stochastic safety verification using barrier certificates," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1415–1428, August 2007.

[15] K. Margellos and J. Lygeros, "Hamilton-Jacobi formulation for reach-avoid differential games," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1849–1861, 2011.

[16] S. Summers and J. Lygeros, "Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem," *Automatica*, vol. 46, no. 12, pp. 1951 – 1961, 2010.

[17] P. M. Esfahani, D. Chatterjee, and J. Lygeros, "The stochastic reach-avoid problem and set characterization for diffusions," *Automatica*, vol. 70, pp. 43 – 56, 2016.

[18] A. Brandejsky, B. de Saporta, and F. Dufur, "Numerical methods for the exit time of a piecewise-deterministic Markov process," *Advances in Applied Probability*, vol. 44, no. 1, pp. 196–225, 2012.

[19] P. E. Kloeden, E. Platen, and H. Schurz, *Numerical solution of SDE through computer experiments*, ser. Universitext. Springer-Verlag, Berlin, 1994.

[20] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*, ser. Imperial College Press Optimization Series. Imperial Collage Press, 2010, vol. 1.

[21] E. de Klerk, M. Laurent, and Z. Sun, "Convergence analysis for Lasserre's measure-based hierarchy of upper bounds for polynomial optimization," *Mathematical Programming*, vol. 162, no. 1, pp. 363–392, Mar 2017.

[22] J. B. Lasserre, "Convergent SDP-relaxations in polynomial optimization with sparsity," *SIAM J. Optim.*, vol. 17, no. 3, pp. 822–843, 2006.
[23] A. Klenke, *Probability theory : a comprehensive course*. London: Springer, 2008.
[24] S. N. Ethier and T. G. Kurtz, *Markov processes – characterization and convergence*, ser. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons Inc., 1986.
[25] J. Jacod, "Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, vol. 31, pp. 235–253, 1974/75.
[26] A. Crudu, A. Debussche, A. Muller, and O. Radulescu, "Convergence of stochastic gene networks to hybrid piecewise deterministic processes," *Ann. Appl. Probab.*, vol. 22, no. 5, pp. 1822–1859, 2012.
[27] L. Arnold, *Stochastic differential equations: theory and applications*. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974, translated from the German.
[28] R. Wisniewski, C. Sloth, M. Bujorianu, and N. Piterman, "Safety verification of piecewise-deterministic Markov processes," in *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*, New York, NY, USA, 2016, pp. 257–266.
[29] M. Bouissou, H. Elmqvist, M. Otter, and A. Benveniste, "Efficient Monte Carlo simulation of stochastic hybrid systems," in *Proceedings of the 10th International Modelica Conference*, 2014, pp. 715–725.

Rafael Wisniewski is a professor and head of the Section of Automation and Control, Department of Electronic Systems, Aalborg University. He received his Ph.D. in Electrical Engineering in 1997, and Ph.D. in Mathematics in 2005. In 2007–2008, he was a control specialist at Danfoss A/S. His research interest is in system theory, particularly in hybrid systems.

Manuela Bujorianu is Research Fellow with Maritime Safety Research Centre, Department of Naval Architecture, University of Strathclyde. Before she was affiliated with several universities: Leicester, Warwick, Manchester, Twente, Cambridge and Stirling. Manuela's research is at the border between computer science, applied mathematics, and systems engineering bridged by probability theory and statistics. She has single authored a monograph at Springer Verlag.

Christoffer Sloth is Associate Professor at SDU Robotics, University of Southern Denmark. Previously, he was Associate Professor at Aalborg University, Denmark. He received his PhD degree from Department of Computer Science at Aalborg University in 2012. His research covers safety verification, polynomial optimization and control of robots.