Reachability Analysis for Controlled Discrete Time Stochastic Hybrid Systems

Saurabh Amin[1], Alessandro Abate[1], Maria Prandini[2], John Lygeros[3], and Shankar Sastry[1]

[1] University of California at Berkeley - Berkeley, USA
{saurabh, aabate, sastry}@eecs.berkeley.edu
[2] Politecnico di Milano - Milano, Italy
prandini@elet.polimi.it
[3] University of Patras - Patras, Greece
lygeros@ee.upatras.gr

Abstract. A model for discrete time stochastic hybrid systems whose evolution can be influenced by some control input is proposed in this paper. With reference to the introduced class of systems, a methodology for probabilistic reachability analysis is developed that is relevant to safety verification. This methodology is based on the interpretation of the safety verification problem as an optimal control problem for a certain controlled Markov process. In particular, this allows to characterize through some optimal cost function the set of initial conditions for the system such that safety is guaranteed with sufficiently high probability. The proposed methodology is applied to the problem of regulating the average temperature in a room by a thermostat controlling a heater.

1 Introduction

Engineering systems like air traffic control systems or infrastructure networks, and natural systems like biological networks exhibit complex behaviors which can often be naturally described by hybrid dynamical models– systems with interacting discrete and continuous dynamics. In many situations the system dynamics are uncertain, and the evolution of the discrete and continuous dynamics as well as the interactions between them are of stochastic nature.

An important problem in hybrid systems theory is that of reachability analysis. In general terms, a reachability analysis problem consists in evaluating if a given system will reach a certain set during some time horizon, starting from some set of initial conditions. This problem arises, for instance, in connection with those safety verification problems where the unsafe conditions for the system can be characterized in terms of its state entering some unsafe set: if the state of the system cannot enter the unsafe set, then the system is declared to be "safe". In a stochastic setting, the safety verification problem can be formulated as that of estimating the probability that the state of the system remains outside the unsafe set for a given time horizon. If the evolution of the state can be influenced by some control input, the problem becomes verifying if it is possible to keep the state of the system outside the unsafe set with sufficiently high probability by selecting a suitable control input.

João Hespanha and A. Tiwari (Eds.): HSCC 2006, LNCS 3927, pp. 49–63, 2006.

Reachability analysis for stochastic hybrid systems has been a recent focus of research, e.g., in [1, 2, 3, 4]. Most of the approaches consider the problem of reachability analysis for continuous time stochastic hybrid systems (CTSHS), wherein the effect of control actions is not directly taken into account. The theory of CTSHS, developed for instance in [5, 6, 7], is used in [1] to address the theoretical issues regarding measurability of the reachability events. On the computational side, a stochastic approximation method is used in [2, 4] to compute the probability of entering into the unsafe set (reach probability). More recently, in [3], certain functions of the state of the system known as barrier certificates are used to compute an upper bound on the reach probability. In the discrete time framework, [8] computes the reach probability using randomized algorithms.

This study adopts the discrete time setting in order to gain a deeper understanding of the theoretical and computational issues associated with the reachability analysis of stochastic hybrid systems. The present work extends the above mentioned approaches to controlled systems, by developing a methodology to compute the maximum probability of remaining in a safe set for a discrete time stochastic hybrid system (DTSHS) whose dynamics is affected by a control input. The approach is based on formulating the reachability analysis problem as an optimal control problem. The maximum probability of remaining in a safe set for a certain time horizon can then be computed by dynamic programming. In addition, the optimal value function obtained through the dynamic programming approach directly enables one to compute the maximal safe set for a specified threshold probability, which is the largest set of all initial conditions such that the probability of remaining in the safe set during a certain time horizon is greater than or equal to the threshold probability.

The paper is organized as follows: Section 2 introduces a model for a DTSHS. This model is inspired by the stochastic hybrid systems models previously introduced in [5, 6, 7] in continuous time. An equivalent representation of the DTSHS in the form of a controlled Markov process is derived. In Section 3, the notion of stochastic reachability for a DTSHS is introduced. The problem of determining probabilistic maximal safe sets for a DTSHS is formulated as a stochastic reachability analysis problem, which can be solved by dynamic programming. The representation of the DTSHS as a controlled Markov process is useful in this respect. In Section 4 we apply the proposed methodology to the problem of regulating the temperature of a room by a thermostat that controls a heater. Concluding remarks are drawn in Section 5.

2 Discrete Time Stochastic Hybrid System

In this section, we introduce a definition of discrete time stochastic hybrid system (DTSHS). This definition is inspired by the continuous time stochastic hybrid system (CTSHS) model described in [9].

The hybrid state of the DTSHS is characterized by a discrete and a continuous component. The discrete state component takes values in a finite set $\mathcal{Q}$. In each mode $q \in \mathcal{Q}$, the continuous state component takes values in the Euclidean space

$\mathbb{R}^{n(q)}$, whose dimension is determined by the map $n : \mathcal{Q} \to \mathbb{N}$. Thus the hybrid state space is $\mathcal{S} := \cup_{q\in\mathcal{Q}}\{q\} \times \mathbb{R}^{n(q)}$. Let $\mathcal{B}(\mathcal{S})$ be the σ-field generated by the subsets of $\mathcal{S}$ of the form $\cup_q\{q\} \times A_q$, where A_q is a Borel set in $\mathbb{R}^{n(q)}$. It can be shown (see [5, page 58]) that $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$ is a Borel space.

The continuous state evolves according to a probabilistic law that depends on the discrete state. A transition from one discrete state to another may occur during the continuous state evolution, according to some probabilistic law. This will then cause a modification of the probabilistic law governing the continuous state evolution. A control input can affect both the continuous and discrete probabilistic evolutions. After a transition in the discrete state has occurred, the continuous state is subject to a probabilistic reset that is also influenced by some control input. Following the reference CTSHS model in [9], we distinguish this latter input from the former one. We call them transition input and reset input, respectively.

Definition 1 (DTSHS). *A discrete time stochastic hybrid system (DTSHS) is a tuple* $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$*, where*

- $\mathcal{Q} := \{q_1, q_2, \ldots, q_m\}$*, for some* $m \in \mathbb{N}$*, represents the discrete state space;*
- $n : \mathcal{Q} \to \mathbb{N}$ *assigns to each discrete state value* $q \in \mathcal{Q}$ *the dimension of the continuous state space* $\mathbb{R}^{n(q)}$*. The hybrid state space is then given by* $\mathcal{S} := \cup_{q\in\mathcal{Q}}\{q\} \times \mathbb{R}^{n(q)}$*;*
- $\mathcal{U}$ *is a compact Borel space representing the transition control space;*
- Σ *is a compact Borel space representing the reset control space;*
- $T_x : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times \mathcal{S} \times \mathcal{U} \to [0, 1]$ *is a Borel-measurable stochastic kernel on* $\mathbb{R}^{n(\cdot)}$ *given* $\mathcal{S} \times \mathcal{U}$*, which assigns to each* $s = (q, x) \in \mathcal{S}$ *and* $u \in \mathcal{U}$ *a probability measure on the Borel space* $(\mathbb{R}^{n(q)}, \mathcal{B}(\mathbb{R}^{n(q)}))$*:* $T_x(dx|(q, x), u)$*;*
- $T_q : \mathcal{Q} \times \mathcal{S} \times \mathcal{U} \to [0, 1]$ *is a discrete stochastic kernel on* $\mathcal{Q}$ *given* $\mathcal{S} \times \mathcal{U}$*, which assigns to each* $s \in \mathcal{S}$ *and* $u \in \mathcal{U}$*, a probability distribution over* $\mathcal{Q}$*:* $T_q(q|s, u)$*;*
- $R : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times \mathcal{S} \times \Sigma \times \mathcal{Q} \to [0, 1]$ *is a Borel-measurable stochastic kernel on* $\mathbb{R}^{n(\cdot)}$ *given* $\mathcal{S} \times \Sigma \times \mathcal{Q}$*, that assigns to each* $s = (q, x) \in \mathcal{S}$*,* $\sigma \in \Sigma$*, and* $q' \in \mathcal{Q}$*, a probability measure on the Borel space* $(\mathbb{R}^{n(q')}, \mathcal{B}(\mathbb{R}^{n(q')}))$*:* $R(dx|(q, x), \sigma, q')$*.* □

In order to define the semantics of a DTSHS, we need first to specify how the system is initialized and how the reset and transition inputs are selected. The system initialization can be specified through some probability measure $\pi : \mathcal{B}(\mathcal{S}) \to [0, 1]$ on the Borel space $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$. When the initial state of the system is $s \in \mathcal{S}$, then, the probability measure π is concentrated at $\{s\}$. As for the choice of the reset and transition inputs, we need to specify which is the rule to determine their values at every time step during the DTSHS evolution (control policy). Here, we consider a DTSHS evolving over a finite horizon $[0, N]$ $(N < \infty)$. If the values for the control inputs at each time $k \in [0, N)$ are determined based on the values taken by the past inputs and the state up to the current time k, then the policy is said to be a feedback policy.

Definition 2 (Feedback policy). *Let $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ be a DTSHS. A feedback policy μ for $\mathcal{H}$ is a sequence $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1})$ of universally measurable maps $\mu_k : \mathcal{S} \times (\mathcal{S} \times \mathcal{U} \times \Sigma)^k \to \mathcal{U} \times \Sigma$, $k = 0, 1, \ldots, N-1$. We denote the set of feedback policies as $\mathcal{M}$.* □

Definition 3 (Execution). *Consider a DTSHS $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$. A stochastic process $\{\mathbf{s}(k) = (\mathbf{q}(k), \mathbf{x}(k)),\ k \in [0, N]\}$ with values in $\mathcal{S} = \cup_{q \in \mathcal{Q}} \{q\} \times \mathbb{R}^{n(q)}$ is an execution of $\mathcal{H}$ associated with a policy $\mu \in \mathcal{M}$ and an initial distribution π if its sample paths are obtained according to the following algorithm, where all the random extractions involved are independent:*

DTSHS algorithm:

Extract from $\mathcal{S}$ a value $s_0 = (q_0, x_0)$ for the random variable $\mathbf{s}(0) = (\mathbf{q}(0), \mathbf{x}(0))$ according to π;

set k=0

while $k < N$ do

set $(u_k, \sigma_k) = \mu_k(s_k, s_{k-1}, u_{k-1}, \sigma_{k-1}, \ldots)$;

extract from Q a value q_{k+1} for the random variable $\mathbf{q}(k+1)$ according to $T_q(\cdot\,|(q_k, x_k), u_k)$;

if $q_{k+1} = q_k$, then

extract from $\mathbb{R}^{n(q_{k+1})}$ a value x_{k+1} for $\mathbf{x}(k+1)$ according to $T_x(\cdot\,|(q_k, x_k), u_k)$

else

extract from $\mathbb{R}^{n(q_{k+1})}$ a value x_{k+1} for $\mathbf{x}(k+1)$ according to $R(\cdot\,|(q_k, x_k), \sigma_k, q_{k+1})$

set $s_{k+1} = (q_{k+1}, x_{k+1})$

$k \to k+1$

end □

If the values for the control inputs are determined only based on the value taken by the state at the current time step, i.e., $(u_k, \sigma_k) = \mu_k(s_k)$, then the policy is said to be a Markov policy.

Definition 4 (Markov Policy). *Consider a DTSHS $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$. A Markov policy μ for $\mathcal{H}$ is a sequence $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1})$ of universally measurable maps $\mu_k : \mathcal{S} \to \mathcal{U} \times \Sigma$, $k = 0, 1, \ldots, N-1$. We denote the set of Markov policies as $\mathcal{M}_m$.*

Note that Markov policies are a subset of the feedback policies: $\mathcal{M}_m \subseteq \mathcal{M}$.

Remark 1. It is worth noticing that the map T_q can model both the spontaneous transitions that might occur during the continuous state evolution, and the forced transitions that must occur when the continuous state exits some prescribed set.

As for spontaneous transitions, if at some hybrid state $(q, x) \in \mathcal{S}$ a transition to the discrete state q' is allowed by the control input $u \in \mathcal{U}$, then this is modeled by $T_q(q'|(q, x), u) > 0$. T_q also encodes a possible delay in the actual occurrence of a transition: if $T_q(q'|(q, x), u) = 1$, then the transition must occur, the smaller is $T_q(q'|(q, x), u)$, the more likely is that the transition will be postponed to a later time.

The invariant set $Dom(q)$ of a discrete state $q \in \mathcal{Q}$, namely the set of all the admissible values for the continuous state within mode q, can be expressed in terms of T_q by forcing $T_q(q|(q, x), u)$ to be zero irrespectively of the value of the control input u in $\mathcal{U}$, for all the continuous state values $x \in \mathbb{R}^{n(q)}$ outside $Dom(q)$. Thus $Dom(q) := \mathbb{R}^{n(q)} \setminus \{x \in \mathbb{R}^{n(q)} : T_q(q|(q, x), u) = 0, \forall u \in \mathcal{U}\}$. □

Define the stochastic kernel $\tau_x : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times \mathcal{S} \times \mathcal{U} \times \Sigma \times \mathcal{Q} \to [0, 1]$ on $\mathbb{R}^{n(\cdot)}$ given $\mathcal{S} \times \mathcal{U} \times \Sigma \times \mathcal{Q}$, which assigns to each $s = (q, x) \in \mathcal{S}$, $u \in \mathcal{U}$, $\sigma \in \Sigma$ and $q' \in \mathcal{Q}$ a probability measure on the Borel space $(\mathbb{R}^{n(q')}, \mathcal{B}(\mathbb{R}^{n(q')}))$ as follows:

$$\tau_x(dx'\,|(q, x), u, \sigma, q') = \begin{cases} T_x(dx'|(q, x), u), & \text{if } q' = q \\ R(dx'|(q, x), \sigma, q'), & \text{if } q' \neq q. \end{cases}$$

In the DTSHS algorithm, τ_x is used to extract a value for the continuous state at time $k + 1$ given the values taken by the hybrid state and the control inputs at time k, and the value extracted for the discrete state at time $k + 1$.

Based on τ_x we can define the Borel-measurable stochastic kernel $T_s : \mathcal{B}(\mathcal{S}) \times \mathcal{S} \times \mathcal{U} \times \Sigma \to [0, 1]$ on $\mathcal{S}$ given $\mathcal{S} \times \mathcal{U} \times \Sigma$, which assigns to each $s = (q, x) \in \mathcal{S}$, $(u, \sigma) \in \mathcal{U} \times \Sigma$ a probability measure on the Borel space $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$ as follows:

$$T_s(ds'\,|s, (u, \sigma)) = \tau_x(dx'\,|s, u, \sigma, q')T_q(q'|s, u), \tag{1}$$

$s, s' = (q', x') \in \mathcal{S}$, $(u, \sigma) \in \mathcal{U} \times \Sigma$. Then, the DTSHS algorithm can be rewritten in a more compact form as:

extract from $\mathcal{S}$ *a value* s_0 *for the random variable* $\mathbf{s}(0)$ *according to* π*;*

set k*=0*

while $k < N$ *do*

set $(u_k, \sigma_k) = \mu_k(s_k, s_{k-1}, u_{k-1}, \sigma_{k-1}, \dots)$*;*

extract from $\mathcal{S}$ *a value* s_{k+1} *for* $\mathbf{s}(k + 1)$ *according to* $T_s(\cdot\,|s_k, (u_k, \sigma_k))$*;*

$k \to k + 1$

end □

This shows that a DTSHS $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ can be described as a controlled Markov process with state space $\mathcal{S} = \cup_{q \in \mathcal{Q}} \{q\} \times \mathbb{R}^{n(q)}$, control space $\mathcal{A} := \mathcal{U} \times \Sigma$, and controlled transition probability function $T_s : \mathcal{B}(\mathcal{S}) \times \mathcal{S} \times \mathcal{A} \to [0, 1]$ defined in (1). This will be referred to in the following as "embedded controlled Markov process" (see, e.g., [10] for an extensive treatment on controlled Markov processes).

As a consequence of this representation of $\mathcal{H}$, the execution $\{\mathbf{s}(k) = (\mathbf{q}(k), \mathbf{x}(k)), k \in [0, N]\}$ associated with $\mu \in \mathcal{M}$ and π is a stochastic process defined on the canonical sample space $\Omega = \mathcal{S}^N$, endowed with its product topology $\mathcal{B}(\Omega)$, with probability measure P_π^μ uniquely defined by the transition kernel T_s, the policy $\mu \in \mathcal{M}$, and the initial probability measure π (see [11, Proposition 7.45]). When π is concentrated at $\{s\}$, $s \in \mathcal{S}$, we shall write simply P_s^μ. From the embedded Markov process representation of a DTSHS it also follows that the execution of a DTSHS associated with a Markov policy μ and an initial condition π is a Markov process. In the sequel, only Markovian policies will be considered.

Example 1 (The thermostat). Consider the problem of regulating the temperature of a room by a thermostat that can switch a heater on and off.

The state of the controlled system is naturally described as a hybrid state. The discrete state component is represented by the heater being in either the "on" or the "off" condition. The continuous state component is represented by the average temperature of the room.

We next show how the controlled system can be described through a DTSHS model $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$. We then formulate the temperature regulation problem with reference to this model.

Concerning the state space of the DTSHS, the discrete component of the hybrid state space is $\mathcal{Q} = \{\texttt{ON}, \texttt{OFF}\}$, whereas $n : \mathcal{Q} \to \mathbb{N}$ defining the dimension of the continuous component of the hybrid state space is the constant map $n(q) = 1, \forall q \in \mathcal{Q}$. We assume that the heater can be turned on or off, and that this is the only available control on the system. We then define $\Sigma = \emptyset$ and $\mathcal{U} = \{0, 1\}$ with the understanding that "1" means that a switching command is issued, "0" that no switching command is issued. Regarding the continuous state evolution, in the stochastic model proposed in [12], the average temperature of the room evolves according to the following stochastic differential equations (SDEs)

$$d\mathbf{x}(t) = \begin{cases} -\frac{a}{C}(\mathbf{x}(t) - x_a)dt + \frac{1}{C}d\mathbf{w}(t), & \text{if the heater is off} \\ -\frac{a}{C}(\mathbf{x}(t) - x_a)dt + \frac{r}{C}dt + \frac{1}{C}d\mathbf{w}(t), & \text{if the heater is on,} \end{cases} \tag{2}$$

where a is the average heat loss rate; C is the average thermal capacity of the room; x_a is the ambient temperature (assumed to be constant); r is the rate of heat gain supplied by the heater; $\mathbf{w}(t)$ is a standard Wiener process modeling the noise affecting the temperature evolution. By applying the constant-step Euler-Maruyama discretization scheme [13] to the SDEs in (2), with time step Δt, we obtain the stochastic difference equation

$$\mathbf{x}(k+1) = \begin{cases} \mathbf{x}(k) - \frac{a}{C}(\mathbf{x}(k) - x_a)\Delta t + \mathbf{n}(k), & \text{if the heater is off} \\ \mathbf{x}(k) - \frac{a}{C}(\mathbf{x}(k) - x_a)\Delta t + \frac{r}{C}\Delta t + \mathbf{n}(k) & \text{if the heater is on,} \end{cases} \tag{3}$$

where $\{\mathbf{n}(k), k \geq 0\}$ is a sequence of i.i.d. Gaussian random variables with zero mean and variance $\nu^2 := \frac{1}{C^2}\Delta t$.

Let $\mathcal{N}(\cdot; m, \sigma^2)$ denote the probability measure over $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ associated with a Gaussian density function with mean m and variance σ^2. Then, the continuous transition kernel T_x implicitly defined in (3) can be expressed as follows:

$$T_x(\cdot\,|(q,x),u) = \begin{cases} \mathcal{N}(\cdot; x - \frac{a}{C}(x - x_a)\Delta t, \nu^2), & q = \texttt{OFF} \\ \mathcal{N}(\cdot; x - \frac{a}{C}(x - x_a)\Delta t + \frac{r}{C}\Delta t, \nu^2), & q = \texttt{ON} \end{cases} \tag{4}$$

Note that the evolution of the temperature within each mode is uncontrolled and so the continuous transition kernel T_x does not depend on the value u of the transition control input.

We assume that it takes some (random) time for the heater to actually switch between its two operating conditions, after a switching command has been issued. This is modeled by defining the discrete transition kernel T_q as follows

$$T_q(q'|(q,x),0) = \begin{cases} 1, & q' = q \\ 0, & q' \neq q \end{cases}$$

$$T_q(q'|(q,x),1) = \begin{cases} \alpha, & q' = \texttt{OFF},\ q = \texttt{ON} \\ 1-\alpha, & q' = q = \texttt{ON} \\ \beta, & q' = \texttt{ON},\ q = \texttt{OFF} \\ 1-\beta, & q' = q = \texttt{OFF} \end{cases} \tag{5}$$

$\forall x \in \mathbb{R}$, where $\alpha \in [0,1]$ represents the probability of switching from the ON to the OFF mode in one time-step. Similarly for $\beta \in [0,1]$.

We assume that the actual switching between the two operating conditions of the heater takes a time step. During this time step the temperature keeps evolving according to the dynamics referring to the starting condition. This is modeled by defining the reset kernel as follows

$$R(\cdot\,|(q,x),q') = \begin{cases} \mathcal{N}(\cdot; x - \frac{a}{C}(x - x_a)\Delta t, \nu^2), & q = \texttt{OFF} \\ \mathcal{N}(\cdot; x - \frac{a}{C}(x - x_a)\Delta t + \frac{r}{C}\Delta t, \nu^2), & q = \texttt{ON}. \end{cases} \tag{6}$$

Let $\bar{x}^-$, $\bar{x}^+ \in \mathbb{R}$, with $\bar{x}^- < \bar{x}^+$. Consider the (stationary) Markov policy $\mu_k : \mathcal{S} \to \mathcal{U}$ defined by

$$\mu_k((q,x)) = \begin{cases} 1, & q = \texttt{ON},\ x \geq \bar{x}^+ \text{ or } q = \texttt{OFF},\ x \leq \bar{x}^- \\ 0, & q = \texttt{ON},\ x < \bar{x}^+ \text{ or } q = \texttt{OFF},\ x > \bar{x}^- \end{cases}$$

that switches the heater on when the temperature drops below $\bar{x}^-$ and off when the temperature goes beyond $\bar{x}^+$.

Suppose that initially the heater is off and the temperature is uniformly distributed in the interval between $\bar{x}^-$ and $\bar{x}^+$, independently of the noise process affecting its evolution. In Figure 1, we report some sample paths of the execution of the DTSHS associated with this policy and initial condition. We plot only the continuous state realizations. The temperature is measured in Fahrenheit degrees ($^\circ F$) and the time in minutes (min). The time horizon N is taken to be 600 min. The discretization time step Δt is chosen to be 1 min. The parameters in equations (4) and (6) are assigned the following values: $x_a = 10.5^\circ F$, $a/C = 0.1\ min^{-1}$, $r/C = 10^\circ F/min$, and $\nu = 1^\circ F$. The switching probabilities

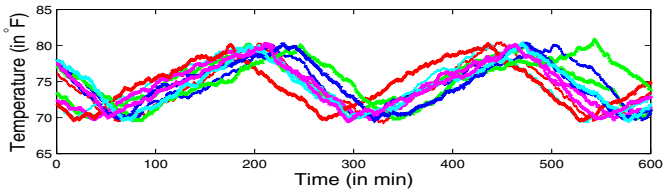

Fig. 1. Sample paths of the temperature for the execution corresponding to a Markov policy switching the heater on/off when the temperature drops below $70°F$/goes above $80°F$, starting with heater off and temperature uniformly distributed on $[70, 80]°F$

α and β in equation (5) are both chosen to be equal to 0.8. Finally, $\bar{x}^-$ and $\bar{x}^+$ are set equal to $70°F$ and $80°F$, respectively.

Note that some of the sample paths exit the set $[70, 80]°F$. This is due partly to the delay in turning the heater on/off and partly to the noise entering the system. If the objective is keeping the temperature within the set $[70, 80]$°F, more effective control policies can be found. In the following section we consider the problem of determining those initial conditions for the system such that it is possible to keep the temperature of the room within prescribed limits over a certain time horizon $[0, N]$, by appropriately acting on the only available control input. Due to the stochastic nature of the controlled system, we relax our requirement to that of keeping the temperature within prescribed limits over $[0, N]$ with sufficiently high probability. We shall see how this problem can be formulated as a stochastic reachability analysis problem. □

3 Stochastic Reachability

We consider the issue of verifying if it is possible to maintain the state of a stochastic hybrid system outside some unsafe set with sufficiently high probability, by choosing an appropriate control policy. This problem can be reinterpreted as a stochastic reachability analysis problem.

With reference to the introduced stochastic hybrid model $\mathcal{H}$, for a given Markov policy $\mu \in \mathcal{M}_m$ and initial state distribution π, a reachability analysis problem consists in determining the probability that the execution associated with the policy μ and initialization π will enter a Borel set $A \in \mathcal{B}(S)$ during the time horizon $[0, N]$:

$$\mathcal{P}_\pi^\mu(A) := P_\pi^\mu(\mathbf{s}(k) \in A \text{ for some } k \in [0, N]). \tag{7}$$

If π is concentrated at $\{s\}$, $s \in \mathcal{S}$, then this is the probability of entering A starting from s, which we denote by $\mathcal{P}_s^\mu(A)$.

Suppose that A represents an unsafe set for $\mathcal{H}$. Different initial conditions are characterized by a different probability of entering A: if the system starts from an initial condition that corresponds to a probability $\epsilon \in (0, 1)$ of entering the unsafe set A, then the system is said to be "safe with probability $1 - \epsilon$". It is

then possible to define sets of initial conditions corresponding to different safety levels, that is sets of states such that the value for the probability of entering the unsafe set starting from them is smaller than or equal to a given value ϵ.

The set of initial conditions that guarantees a safety level $1-\epsilon$, when the control policy $\mu \in \mathcal{M}_m$ is assigned,

$$S^{\mu}(\epsilon) = \{s \in \mathcal{S} : \mathcal{P}_s^{\mu}(A) \leq \epsilon\} \tag{8}$$

is referred to as *probabilistic safe set* with safety level $1-\epsilon$. If the control policy can be selected so as to minimize the probability of entering A, then

$$S^{\star}(\epsilon) = \{s \in \mathcal{S} : \inf_{\mu \in \mathcal{M}_m} \mathcal{P}_s^{\mu}(A) \leq \epsilon\}. \tag{9}$$

is the *maximal probabilistic safe set* with safety level $1-\epsilon$. By comparing the expressions for $S^{\mu}(\epsilon)$ and $S^{\star}(\epsilon)$, it is in fact clear that $S^{\mu}(\epsilon) \subseteq S^{\star}(\epsilon)$, for each $\mu \in \mathcal{M}_m$, $\epsilon \in (0,1)$.

In the rest of the section, we show that (i) the problem of computing $\mathcal{P}_s^{\mu}(A)$ and $S^{\mu}(\epsilon)$ for $\mu \in \mathcal{M}_m$ can be solved by using a backward iterative procedure; and (ii) the problem of computing $S^{\star}(\epsilon)$ can be reduced to an optimal control problem. This, in turn, can be solved by dynamic programming. These results are obtained based on the representation of $\mathcal{P}_{\pi}^{\mu}(A)$ as a multiplicative cost function.

The probability $\mathcal{P}_{\pi}^{\mu}(A)$ defined in (7) can be expressed as $\mathcal{P}_{\pi}^{\mu}(A) = 1 - p_{\pi}^{\mu}(\bar{A})$, where $\bar{A}$ denotes the complement of A in $\mathcal{S}$ and $p_{\pi}^{\mu}(\bar{A}) := P_{\pi}^{\mu}(\mathbf{s}(k) \in \bar{A} \text{ for all } k \in [0,N])$. Let $\mathbf{1}_C : \mathcal{S} \to \{0,1\}$ denote the indicator function of a set $C \subseteq \mathcal{S}$: $\mathbf{1}_C(s) = 1$, if $s \in C$, and 0, if $s \notin C$. Observe that

$$\prod_{k=0}^{N} \mathbf{1}_{\bar{A}}(s_k) = \begin{cases} 1, & \text{if } s_k \in \bar{A} \text{ for all } k \in [0,N] \\ 0, & \text{otherwise,} \end{cases}$$

where $s_k \in \mathcal{S}$, $k \in [0,N]$. Then,

$$p_{\pi}^{\mu}(\bar{A}) = P_{\pi}^{\mu}(\prod_{k=0}^{N} \mathbf{1}_{\bar{A}}(\mathbf{s}(k)) = 1) = E_{\pi}^{\mu}[\prod_{k=0}^{N} \mathbf{1}_{\bar{A}}(\mathbf{s}(k))]. \tag{10}$$

From this expression it follows that

$$p_{\pi}^{\mu}(\bar{A}) = \int_{\mathcal{S}} E_{\pi}^{\mu}\big[\prod_{k=0}^{N} \mathbf{1}_{\bar{A}}(\mathbf{s}(k)) \big| \, s(0) = s\big] \pi(ds), \tag{11}$$

where the conditional mean $E_{\pi}^{\mu}[\prod_{k=0}^{N} \mathbf{1}_{\bar{A}}(\mathbf{s}(k)) | \, s(0) = s]$ is well defined over the support of the probability measure π representing the distribution of $\mathbf{s}(0)$.

3.1 Backward Reachability Computations

We next show how it is possible to compute $p_{\pi}^{\mu}(\bar{A})$ through a backward iterative procedure for a given Markov policy $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1}) \in \mathcal{M}_m$, with μ_k :

$\mathcal{S} \to \mathcal{U} \times \Sigma$, $k = 0, 1, \ldots, N-1$. For each $k \in [0, N]$, define the map $V_k^\mu : \mathcal{S} \to [0,1]$ as follows

$$V_k^\mu(s) := \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}^{N-k}} \prod_{l=k+1}^{N} \mathbf{1}_{\bar{A}}(s_l) \prod_{h=k+1}^{N-1} T_s(ds_{h+1}|s_h, \mu_h(s_h)) T_s(ds_{k+1}|s, \mu_k(s)), \tag{12}$$

$\forall s \in \mathcal{S}$, where T_s is the controlled transition function of the embedded controlled Markov process, and $\int_{\mathcal{S}^0}(\ldots) = 1$. If s belongs to the support of π, then, $E_\pi^\mu\big[\prod_{l=k}^{N} \mathbf{1}_{\bar{A}}(\mathbf{s}(l))|\, \mathbf{s}(k) = s\big]$ is well-defined and equal to the right-hand-side of (12), so that

$$V_k^\mu(s) = E_\pi^\mu\big[\prod_{l=k}^{N} \mathbf{1}_{\bar{A}}(\mathbf{s}(l))|\, \mathbf{s}(k) = s\big] \tag{13}$$

denotes the probability of remaining outside A during the (residual) time horizon $[k, N]$ starting from s at time k, under policy μ applied from π.

By (11) and (13), $p_\pi^\mu(\bar{A})$ can be expressed as $p_\pi^\mu(\bar{A}) = \int_{\mathcal{S}} V_0^\mu(s)\pi(ds)$. If π is concentrated at $\{s\}$, $p_s^\mu(\bar{A}) = V_0^\mu(s)$. Since $\mathcal{P}_s^\mu(A) = 1 - p_s^\mu(\bar{A})$, then the probabilistic safe set with safety level $1-\epsilon$, $\epsilon \in (0,1)$, defined in (8) can be computed as $S^\mu(\epsilon) = \{s \in \mathcal{S} : V_0^\mu(s) \geq 1 - \epsilon\}$.

By a reasoning similar to [14] for additive costs, we prove the following lemma.

Lemma 1. *Fix a Markov policy μ. The maps $V_k^\mu : \mathcal{S} \to [0,1]$, $k = 0, 1 \ldots, N$, can be computed by the backward recursion:*

$$V_k^\mu(s) = \boldsymbol{1}_{\bar{A}}(s)\Big[T_q(q|s, u_k^\mu(s)) \int_{\mathbb{R}^{n(q)}} V_{k+1}^\mu((q,x'))T_x(dx'|s, u_k^\mu(s)) + \sum_{q' \neq q} T_q(q'|s, u_k^\mu(s)) \int_{\mathbb{R}^{n(q')}} V_{k+1}^\mu((q',x'))R(dx'|s, \sigma_k^\mu(s), q')\Big], s = (q,x) \in \mathcal{S},$$

where $\mu_k = (u_k^\mu, \sigma_k^\mu) : \mathcal{S} \to \mathcal{U} \times \Sigma$, initialized with $V_N^\mu(s) = \boldsymbol{1}_{\bar{A}}(s)$, $s \in \mathcal{S}$.

Proof. From definition (12) of V_k^μ, we get that $V_N^\mu(s) = \mathbf{1}_{\bar{A}}(s)$, $s \in \mathcal{S}$. For $k < N$,

$$\begin{aligned} V_k^\mu(s) &= \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}^{N-k}} \prod_{l=k+1}^{N} \mathbf{1}_{\bar{A}}(s_l) \prod_{h=k+1}^{N-1} T_s(ds_{h+1}|s_h, \mu_h(s_h)) T_s(ds_{k+1}|s, \mu_k(s)) \\ &= \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}} \mathbf{1}_{\bar{A}}(s_{k+1}) \Big(\int_{\mathcal{S}^{N-k-1}} \prod_{l=k+2}^{N} \mathbf{1}_{\bar{A}}(s_l) \prod_{h=k+2}^{N-1} T_s(ds_{h+1}|s_h, \mu_h(s_h)) \\ &\quad T_s(ds_{k+2}|s_{k+1}, \mu_{k+1}(s_{k+1}))\Big) T_s(ds_{k+1}|s, \mu_k(s)) \\ &= \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}} V_{k+1}^\mu(s_{k+1}) T_s(ds_{k+1}|s, \mu_k(s)). \end{aligned}$$

Recalling the definition of T_s the thesis immediately follows. □

3.2 Maximal Probabilistic Safe Set Computation

The calculation of the maximal probabilistic safe set $S^\star(\epsilon)$ defined in (9) amounts to finding the infimum over the Markov policies of the probability $\mathcal{P}_s^\mu(A)$ of entering the unsafe set A starting from s, for all s outside A (the probability of entering A starting from $s \in A$ is 1 for any policy). A policy that achieves this infimum is said to be *maximally safe*.

Definition 5 (Maximally safe policy). *Let $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ be a DTSHS, and $A \in \mathcal{B}(\mathcal{S})$ an unsafe set. A policy $\mu^* \in \mathcal{M}_m$ is maximally safe if $\mathcal{P}_s^{\mu^*}(A) = \inf_{\mu \in \mathcal{M}_m} \mathcal{P}_s^\mu(A)$, $\forall s \in \bar{A}$.*

Given that $\mathcal{P}_s^\mu(A) = 1 - p_s^\mu(\bar{A})$, finding the infimum of the probability $\mathcal{P}_s^\mu(A)$ is equivalent to computing the supremum of the probability $p_s^\mu(\bar{A})$ of remaining within the safe set $\bar{A}$. In the following theorem, we describe an algorithm to compute $\sup_{\mu \in \mathcal{M}_m} p_s^\mu(\bar{A})$ and give a condition for the existence of a maximally safe policy. The proof is based on [11, Proposition 11.7].

Theorem 1. *Define the maps $V_k^* : \mathcal{S} \to [0,1]$, $k = 0, 1, \ldots, N$, by the backward recursion:*

$$V_k^*(s) = \sup_{(u,\sigma) \in \mathcal{U} \times \Sigma} \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}} V_{k+1}^*(s_{k+1}) T_s(ds_{k+1} | s, (u, \sigma)),\ s \in \mathcal{S},$$

initialized with $V_N^*(s) = \mathbf{1}_{\bar{A}}(s)$, $s \in \mathcal{S}$.

Then, $V_0^*(s) = \sup_{\mu \in \mathcal{M}_m} p_s^\mu(\bar{A})$ for all $s \in \mathcal{S}$. Moreover, if $U_k(s, \lambda) = \{(u, \sigma) \in \mathcal{U} \times \Sigma | \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}} V_{k+1}^*(s_{k+1}) T_s(ds_{k+1} | s, (u, \sigma)) \leq \lambda\}$ is compact for all $s \in \mathcal{S}$, $\lambda \in \mathbb{R}$, $k \in [0, N-1]$, then there exists a maximally safe policy $\mu^* = (\mu_0^*, \ldots, \mu_{N-1}^*)$, with $\mu_k^* : \mathcal{S} \to \mathcal{U} \times \Sigma$, $k \in [0, N-1]$, given by

$$\mu_k^*(s) = \arg \sup_{(u,\sigma) \in \mathcal{U} \times \Sigma} \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}} V_{k+1}^*(s_{k+1}) T_s(ds_{k+1} | s, (u, \sigma)),\ \forall s \in \mathcal{S}. \quad (14)$$

Proof. Note that we deal with Borel spaces and with Borel measurable stochastic kernels. The one-stage cost function $\mathbf{1}_{\bar{A}}(s)$ is Borel measurable, non negative and bounded for all $s \in \mathcal{S}$. In particular, $V_N^*(s) = \mathbf{1}_{\bar{A}}(s)$ is Borel measurable, hence universally measurable. It can be directly checked that the mapping $H : \mathcal{S} \times \mathcal{U} \times \Sigma \times V \to \mathbb{R}$ defined as $H(s, (u, \sigma), V) = \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}} V(s') T_s(ds' | s, (u, \sigma))$ satisfies the *monotonicity assumption* when applied to universally measurable functions V (cf. [11, Section 6.1]). Then $V_k^*(s) = \sup_{(u,\sigma) \in \mathcal{U} \times \Sigma} H(s, (u, \sigma), V_{k+1}^*)$ is universally measurable for every $k \in [0, N-1]$. The functions $V_k^*(s)$ are also lower semi-analytic. This holds because the product of a lower semi-analytic function by a positive Borel measurable function is lower semi-analytic; furthermore, the integration of a lower semi-analytic function with respect to a stochastic kernel and its supremization with respect to one of its arguments (in this specific instance, the control input) is lower semi-analytic (cf. [11, Propositions 7.30, 7.47 and 7.48]). The preceding measurability arguments provide a solid ground

for the *exact selection assumption* to hold ([11, Section 6.2]), which finally leads to the statement of the theorem by the application of [11, Proposition 11.7]. □

Remark 2. When $\mathcal{U}$ and Σ are finite sets, then the compactness assumption required in the theorem is trivially satisfied.

The maximal probabilistic safe set $S^\star(\epsilon)$ with safety level $1-\epsilon$ defined in (9) can be determined as $S^\star(\epsilon) = \{s \in \mathcal{S} : V_0^*(s) \geq 1-\epsilon\}$.

4 The Thermostat Example

In this section we apply the proposed methodology to the problem of regulating the temperature of a room by a thermostat controlling a heater. We refer to the DTSHS description of the system given in Example 1 of Section 2. The system parameters and time horizon are set equal to the values reported at the end of Example 1. Three safe sets are considered: $\bar{A}_1 = (70, 80)^\circ F$, $\bar{A}_2 = (72, 78)^\circ F$, and $\bar{A}_3 = (74, 76)^\circ F$. The dynamic programming recursion described in Section 3.2 is used to compute maximally safe policies and maximal probabilistic safe sets. The implementation is done in MATLAB. The temperature is discretized into 100 equally spaced values within the safe set.

Figure 2 show the plots of 100 temperature sample paths resulting from sampling the initial temperature from the uniform distribution over the safe sets,

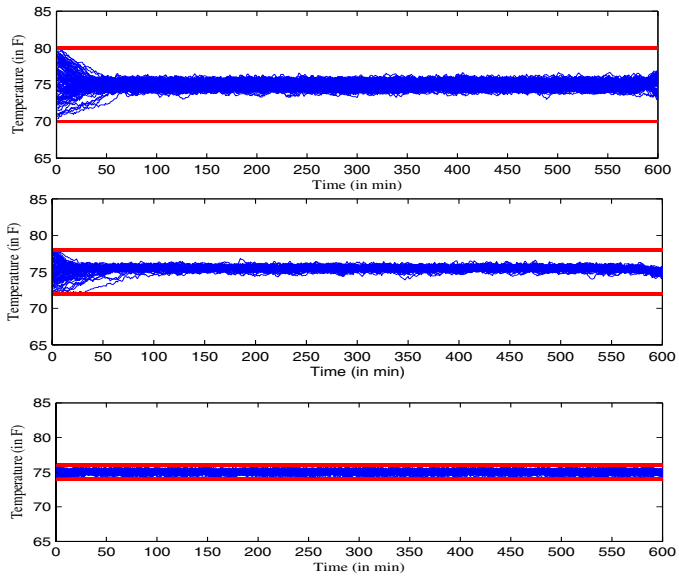

Fig. 2. Sample paths of the temperature for the execution corresponding to maximally safe policies, when the safe set is: $\bar{A}_1$ (top), $\bar{A}_2$ (middle), and $\bar{A}_3$ (bottom)

and using the corresponding maximally safe policy. The initial operating mode is chosen at random between the equiprobable `ON` and `OFF` values.

It can be observed from each of the plots that the maximally safe policy computed by the dynamic programming recursion leads to an optimal behavior in the following sense: regardless of the initial state, most of the temperature sample paths tend toward the middle of the corresponding safe set. As for the $\bar{A}_1$ and $\bar{A}_2$ safe sets, the temperature actually remain confined within the safe set in almost all the sample paths, whereas this is not the case for $\bar{A}_3$. The set $\bar{A}_3$ is too small to enable the control input to counteract the drifts and the randomness in the execution in order to maintain the temperature within the safe set. The maximal probability of remaining in the safe set $p_\pi^{\mu^*}(\bar{A}_i)$ for π uniform over $\mathcal{Q} \times \bar{A}_i$, $i = 1, 2, 3$, is computed. The value is 0.991 for $\bar{A}_1$, 0.978 for $\bar{A}_2$ and 0.802 for $\bar{A}_3$.

The maximal probabilistic safe sets $S^\star(\epsilon)$ corresponding to different safety levels $1 - \epsilon$ are also calculated. The results obtained are reported in Figure 3 with reference to the heater initially off (plot on the left) and on (plot on the right). In all cases, as expected, the maximal probabilistic safe sets get smaller as the required safety level $1 - \epsilon$ grows. When the safe set is $\bar{A}_3$, there is no policy that can guarantee a safety probability greater than about 0.86.

The maximally safe policies at some time instances $k \in [0, 600]$ $\mu_k^* : \mathcal{S} \to \mathcal{U}$ are shown in Figure 4, as a function of the continuous state and discrete state (the red crossed line refers to the `OFF` mode, whereas the blue circled line refers to the `ON` mode). The obtained result is quite intuitive. For example, at time $k = 599$, close to the end of the time horizon, and in the `OFF` mode, the maximally safe policy prescribes to stay in same mode for most of the continuos state values except near the lower boundary of the safe set, in which case it prescribes to change the mode to `ON` since there is a possibility of entering the unsafe set in the residual one-step time horizon. However, at earlier times (for instance, time $k = 1$), the maximally safe policy prescribes to change the mode even for states that are distant from the safe set boundary. Similar comments apply to

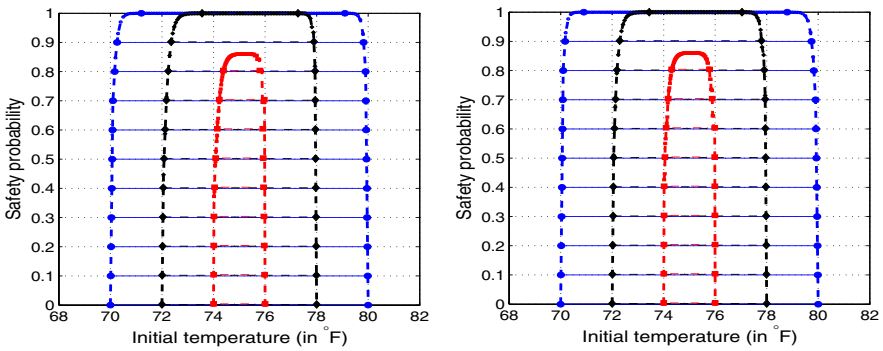

Fig. 3. Maximal probabilistic safe sets: heater initially off (left) and on (right). Blue, black, and red colors refer to cases when the safe sets are $\bar{A}_1$, $\bar{A}_2$, and $\bar{A}_3$, respectively.

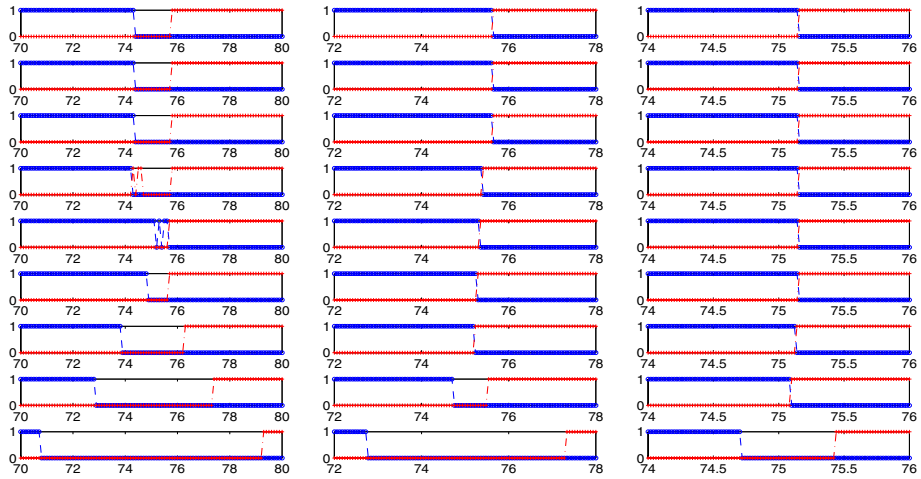

Fig. 4. Maximally safe policy as a function of the temperature at times $k = 1$, 250, 500, 575, 580, 585, 590, 595, and 599 (from top to bottom) for the safe sets $\bar{A}_1$, $\bar{A}_2$, and $\bar{A}_3$ (from left to right). The darker (blue) circled line corresponds to the `OFF` mode and the lighter (red) crossed line corresponds to the `ON` mode.

the `ON` mode. This shows that a maximally safe policy is not stationary. By observing from top to bottom each column of Figure 4, one can see that this non-stationary behavior appears limited to a time interval at the end of the time horizon. Also, by comparing the columns of Figure 4, this time interval gets progressively smaller moving from $\bar{A}_1$ to $\bar{A}_2$ and $\bar{A}_3$.

It is interesting to note the behavior of the maximally safe policy corresponding to the safe set $\bar{A}_1$ at $k = 575$ and $k = 580$. For example, for $k = 580$, the maximally safe policy for the `OFF` mode fluctuates between actions 0 and 1 when the temperature is around $75°F$. This is because the corresponding values taken by the function to be optimized in (14) are almost equal for the two control actions. The results obtained refer to the case of switching probabilities $\alpha = \beta = 0.8$. Different choices of switching probabilities may yield qualitatively different maximally safe policies.

5 Final Remarks

In this paper we proposed a model for controlled discrete time stochastic hybrid systems. With reference to such a model, we described the notion of stochastic reachability, and discussed how the problem of safety verification can be reinterpreted in terms of the introduced stochastic reachability notion. By an appropriate reformulation of the safety verification problem for the stochastic hybrid system as that of determining a feedback policy that optimizes some multiplicative cost function for a certain controlled Markov process, we were able to suggest a solution based on dynamic programming. Temperature regulation of a

room by a heater that can be repeatedly switched on and off was presented as a simple example to illustrate the model capabilities and the reachability analysis methodology.

Further work is needed to extend the current approach to the infinite horizon and partial information cases. The more challenging problem of stochastic reachability analysis for continuous time stochastic hybrid systems is an interesting subject of future research.

References

1. Bujorianu, M.L., Lygeros, J.: Reachability questions in piecewise deterministic Markov processes. In Maler, O., Pnueli, A., eds.: Hybrid Systems: Computation and Control. Lecture Notes in Computer Science 2623. Springer Verlag (2003) 126–140
2. Hu, J., Prandini, M., Sastry, S.: Probabilistic safety analysis in three-dimensional aircraft flight. In: Proc. of the IEEE Conf. on Decision and Control (2003)
3. Prajna, S., Jadbabaie, A., J.Pappas, G.: Stochastic safety verification using barrier certificates. In: Proc. of the IEEE Conf. on Decision and Control (2004)
4. Hu, J., Prandini, M., Sastry, S.: Aircraft conflict prediction in the presence of a spatially correlated wind field. IEEE Trans. on Intelligent Transportation Systems **6**(3) (2005) 326–340
5. Davis, M.H.A.: Markov Models and Optimization. Chapman & Hall, London (1993)
6. Ghosh, M.K., Arapostathasis, A., Marcus, S.I.: Ergodic control of switching diffusions. SIAM Journal of Control and Optimization **35**(6) (1997) 1952–1988
7. Hu, J., Lygeros, J., Sastry, S.: Towards a theory of stochastic hybrid systems. In Lynch, N., Krogh, B., eds.: Hybrid Systems: Computation and Control. Lecture Notes in Computer Science 1790. Springer Verlag (2000) 160–173
8. Lygeros, J., Watkins, O.: Stochastic reachability for discrete time systems: an application to aircraft collision avoidance. In: Proc. of the IEEE Conf. on Decision and Control (2003)
9. Bujorianu, M., Lygeros, J.: General stochastic hybrid systems: Modelling and optimal control. In: Proc. of the IEEE Conf. on Decision and Control. (2004)
10. Puterman, M.: Markov decision processes. John Wiley & Sons, Inc (1994)
11. Bertsekas, D.P., Shreve, S.E.: Stochastic optimal control: the discrete-time case. Athena Scientific (1996)
12. Malhame, R., Chong, C.Y.: Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system. IEEE Transactions on Automatic Control **AC-30**(9) (1985) 854–860
13. Milstein, G.: Numerical Integration of Stochastic Differential Equations. Kluver Academic Press (1995)
14. Kumar, P.R., Varaiya, P.P.: Stochastic Systems: Estimation, Identification, and Adaptive Control. Prentice Hall, Inc., New Jersey (1986)