Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

Probabilistic reachability and safe sets computation for discrete time stochastic hybrid systems

Alessandro Abate†, Saurabh Amin†, Maria Prandini‡, John Lygeros* and Shankar Sastry†

Abstract— **In this work probabilistic reachability for controlled discrete time stochastic hybrid systems is investigated. By a suitable formulation of the reachability problem within a stochastic optimal control framework, two complementary interpretations and their corresponding computational techniques are suggested. The results can be of interest for solving safety analysis and control design problems for stochastic hybrid systems, by the computation of maximal probabilistic safe sets and maximally safe policies. They can also be employed to solve regulation problems through the interpretation of the desired operating region for the system as a "safe set". The described methodology is applied to a simple temperature regulation problem.**

I. Introduction

An important topic in classical control system theory is reachability, which can be also interpreted as the problem of maintaining the state of the system within some prespecified set by selecting a suitable control law [1]. Recently, this problem has attracted even more the interest of the control community due to its application to real-time automation systems such as, for example, air traffic control [2]. Such applications introduce additional complexity to the problem, in that they typically involve the interaction between discrete and continuous dynamic components, often in the presence of uncertainty affecting their evolution. This motivates the study of reachability for controlled stochastic hybrid systems (SHS), which is the topic discussed in this paper.

We consider a discrete time SHS (DTSHS), and address the issue of determining if the state of the DTSHS can be maintained *within* a given "safe" set with sufficiently high probability by applying a suitable control input. As shown in [3] with reference to the finite time horizon case, by adopting an optimal control viewpoint, the problem can be formulated as that of determining the feedback control law that maximizes the probability of the state of the controlled system to evolve within the safe set. Based on the expression of this probability as a *multiplicative cost* function, and restricting the controller class to static state feedback control laws (Markov policies), dynamic programming (DP) can be effectively used to compute *probabilistic maximal safe sets* corresponding to different safety levels. These are the initial states for the system, such that there exists a control law capable of maintaining the state of the system within the safe set with a probability not smaller than a prescribed safety level (see [4], [5] for the corresponding definition in the deterministic case).

In this paper, we investigate the complementary problem of keeping the state of a DTSHS *outside* some prespecified "unsafe" set by selecting a suitable feedback control law. We again formulate the problem as an optimal control problem wherein the objective is now to minimize some *max cost* function. We show that the DP approach is still effective for determining probabilistic maximal safe sets for Markov policies. In fact, the value functions for the max cost case can be expressed in terms of the value functions for the multiplicative cost case, thus formalizing the intuition that the two viewpoints for reachability analysis are complementary to each other.

We characterize the maximally safe control law within the class of Markov policies. We also extend reachability analysis for DTSHS to the infinite time-horizon setting and address the question of convergence of the optimal control law to a stationary policy.

Reachability can also be studied within the framework of *regulation* theory where the aim is to steer the state of the system close to some desired operating condition. This can be achieved by considering a small neighborhood around the desired operating condition, and by solving a reachability problem with a time-varying region that shrinks to that neighborhood as the "safe" set for the system. If the state of the system has then to be maintained within this neighborhood indefinitely, [6], one can split the problem into a finite horizon time-varying reachability problem and a subsequent infinite horizon one. This approach has close connections with control design for *practical stabilization*, [7]. The application of reachability analysis to regulation problems for DTSHS is discussed here with reference to a simple application example, where the problem is to drive the temperature of a room close to some desired value by controlling a heater.

The rest of the paper is organized as follows. In Section II we recall the definition of DTSHS given in [3]. We then describe in Section III the stochastic reachability analysis problem for Markov policies. We address it according to the two complementary viewpoints that lead to a DP solution using a multiplicative and a max cost function. In Section IV, we consider the problem of probabilistic maximal safe set computation, and discuss conditions such that a maximally safe Markov policy exists. A numerical example is given with reference to the problem of temperature regulation. The

Research supported by the European Commission under project HYGEIA, FP6-NEST-004995 and the Network of Excellence HYCON, FP6-IST-511368, by MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca) under the project 'New methods for Identification and Adaptive Control for Industrial Systems', and by the NSF grant CCR-0225610.

† University of California at Berkeley, CA, USA – {aabate,saurabh,sastry}@eecs.berkeley.edu
‡ Politecnico di Milano, Italy – prandini@elet.polimi.it
* ETH Zurich, Switzerland – lygeros@control.ee.ethz.ch

extension to the infinite horizon case is treated in Section V. Finally, some concluding remarks are drawn in Section VI.

II. Discrete time stochastic hybrid system model

In this section, we consider the discrete time stochastic hybrid system (DTSHS) model introduced in [3]. The state of the DTSHS is characterized by a discrete and a continuous component. The continuous state evolves according to a probabilistic law that depends on the value taken by the discrete state. In turn, the discrete state can transition between different values in a set according to some probabilistic law that depends on the value taken by the continuous state. Both the continuous and the discrete probabilistic evolutions can be affected by some control input (transition input). After a transition in the discrete state has occurred, the continuous state is subject to a probabilistic reset that depends on some control input (reset input).

Definition 1: 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 $\mathbb{R}^{n(q)}$.
- $\mathcal{U}$ 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 $T_x(dx|s,u)$ on the Borel space $(\mathbb{R}^{n(q)}, \mathcal{B}(\mathbb{R}^{n(q)}))$.
- $T_q : \mathcal{Q} \times \mathcal{S} \times \mathcal{U} \to [0,1]$ 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 $T_q(q|s,u)$ over $\mathcal{Q}$.
- $R : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times \mathcal{S} \times \Sigma \times \mathcal{Q} \to [0,1]$ is a Borel-measurable stochastic kernel on $\mathbb{R}^{n(\cdot)}$ given $\mathcal{S} \times \Sigma \times \mathcal{Q}$, that assigns to each $s \in \mathcal{S}$, $\sigma \in \Sigma$, and $q' \in \mathcal{Q}$, a probability measure $R(dx|s,\sigma,q')$ on the Borel space $(\mathbb{R}^{n(q')}, \mathcal{B}(\mathbb{R}^{n(q')}))$. □

The hybrid state space is $\mathcal{S} := \cup_{q \in \mathcal{Q}} \{q\} \times \mathbb{R}^{n(q)}$. $\mathcal{B}(\mathcal{S})$ is the σ-field generated by the subsets of $\mathcal{S}$ of the form $\cup_q \{q\} \times A_q$, with A_q denoting a Borel set in $\mathbb{R}^{n(q)}$.

In order to define an execution for a DTSHS we have to specify how the system is initialized and how the control inputs to the system are selected.

The system initialization at time $k = 0$ is specified through some probability measure $\pi : \mathcal{B}(\mathcal{S}) \to [0,1]$ on the Borel space $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$. As for the control input, we next define the notion of feedback policy with reference to a finite time horizon $[0, N]$.

Definition 2 (Feedback policy): A feedback policy for a DTSHS $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ 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 $\mathcal{M}$. □

Let $\tau_x : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times \mathcal{S} \times \mathcal{U} \times \Sigma \times \mathcal{Q} \to [0,1]$ be a stochastic kernel 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:

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

Based on τ_x we can introduce 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 $s = (q,x), s' = (q',x') \in \mathcal{S}$, $(u,\sigma) \in \mathcal{U} \times \Sigma$ a probability measure on the Borel space $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$ as follows:

$$T_s(ds' \,|s, (u,\sigma)) = \tau_x(dx' \,|s, u, \sigma, q') T_q(q'|s,u). \qquad (1)$$

Definition 3 (Execution): An execution for a DTSHS $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ associated with a policy $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1}) \in \mathcal{M}$ and an initial distribution π is a stochastic process $\{\mathbf{s}(k), k \in [0, N]\}$ with values in $\mathcal{S}$ whose sample paths are obtained according to the following algorithm:

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

for $k = 0$ *to* $N-1$

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

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

end □

A DTSHS $\mathcal{H}$ is a controlled Markov process (see, e.g., [8], [9]) with state space $\mathcal{S}$, control space $\mathcal{U} \times \Sigma$, and controlled transition probability function $T_s : \mathcal{B}(\mathcal{S}) \times \mathcal{S} \times \mathcal{U} \times \Sigma \to [0,1]$ defined in (1). Thus, the execution $\{\mathbf{s}(k), k \in [0,N]\}$ associated with $\mu \in \mathcal{M}$ and π is a stochastic process defined on the canonical sample space $\Omega = \mathcal{S}^{N+1}$, 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 [9, Proposition 7.45]).

If $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1}) \in \mathcal{M}$ is such that for any $k \in [0, N-1]$ the values for the control inputs (u_k, σ_k) are determined only based on the value taken by the state at the time step k, then the policy is said to be Markov.

Definition 4 (Markov Policy): A Markov policy for a DTSHS $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ 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 $\mathcal{M}_m$. □

If $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1})$ is a Markov policy, then, the execution of $\mathcal{H}$ associated with μ and a distribution π on $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$ is a time inhomogeneous Markov process with initial state distribution π, and one-step transition kernels $T_s^{\mu_k}(ds'|s) := T_s(ds'|s, \mu_k(s))$, $k = 0, 1, \ldots, N-1$.

In the rest of the paper we shall consider only feedback policies that are Markov.

III. Stochastic reachability

We consider the following reachability problem: Given a stochastic hybrid system $\mathcal{H}$, determine the probability that

the execution associated with some policy $\mu \in \mathcal{M}_m$ and initialization π will enter a Borel set $A \in \mathcal{B}(\mathcal{S})$ during the time horizon $[0, N]$:

$$\mathcal{P}_\pi^\mu(A) := P_\pi^\mu(\mathbf{s}(k) \in A \text{ for some } k \in [0, N]). \quad (2)$$

If π is concentrated at $s \in \mathcal{S}$, i.e. $\pi(ds) = \delta_s(ds)$, then $\mathcal{P}_\pi^\mu(A)$ represents the probability of entering A starting from s. Hence, we denote it by $\mathcal{P}_s^\mu(A)$. When A is an unsafe set for $\mathcal{H}$, by computing $\mathcal{P}_s^\mu(A)$, we evaluate the safety level for system $\mathcal{H}$ when it starts from s and is subject to policy μ.

In this section we show that for a Markov policy μ the problem of computing $\mathcal{P}_s^\mu(A)$ can be solved by using an iterative procedure.

A. Max cost

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

$$\max_{k\in[0,N]} \mathbf{1}_A(s_k) = \begin{cases} 1, & \text{if } s_k \in A \text{ for some } k \in [0, N] \\ 0, & \text{otherwise,} \end{cases}$$

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

The probability $\mathcal{P}_\pi^\mu(A)$ in (2) can then be expressed as

$$\mathcal{P}_\pi^\mu(A) = E_\pi^\mu\big[\max_{k\in[0,N]} \mathbf{1}_A(\mathbf{s}(k))\big]. \quad (3)$$

From this expression it follows that

$$\mathcal{P}_\pi^\mu(A) = \int_{\mathcal{S}} E_\pi^\mu\big[\max_{k\in[0,N]} \mathbf{1}_A(\mathbf{s}(k))|\, s(0) = s\big]\pi(ds), \quad (4)$$

where $E_\pi^\mu[\max_{k\in[0,N]} \mathbf{1}_A(\mathbf{s}(k))|\, s(0) = s]$ is well defined over the support of the distribution π of $\mathbf{s}(0)$.

Consider now a Markov policy $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1}) \in \mathcal{M}_m$, with $\mu_k : \mathcal{S} \to \mathcal{U} \times \Sigma$, $\forall k \in \{0, 1, \ldots, N-1\}$, and an initial distribution π. For each $k \in [0, N]$, introduce the function $W_k^\mu : \mathcal{S} \to [0, 1]$ as follows

$$W_k^\mu(s) := \mathbf{1}_{\bar{A}}(s) \int_{\mathcal{S}^{k+1}} \max\{0, \mathbf{1}_A(s_h),\, N-k+1 \le h \le N\} \prod_{h=N-k}^{N-1} T_s^{\mu_h}(ds_{h+1}|s_h)\delta_s(ds_{N-k}) + \mathbf{1}_A(s),\ s \in \mathcal{S}, \quad (5)$$

where, for $k = 0, W_0^\mu(s) = \mathbf{1}_A(s)$, and $\bar{A}$ denotes the complement of A in $\mathcal{S}$, $\bar{A} = \mathcal{S} \setminus A$. It is easily seen that the right-hand-side of (5) can be rewritten as

$$\int_{\mathcal{S}^{k+1}} \max_{h\in[N-k,N]} \mathbf{1}_A(s_h) \prod_{h=N-k}^{N-1} T_s^{\mu_h}(ds_{h+1}|s_h)\delta_s(ds_{N-k})$$

so that, for those $s \in \mathcal{S}$ belonging to the support of the distribution of $\mathbf{s}(N-k)$,

$$W_k^\mu(s) = E_\pi^\mu\big[\max_{l\in[N-k,N]} \mathbf{1}_{\bar{A}}(\mathbf{s}(l))|\, \mathbf{s}(N-k) = s\big]. \quad (6)$$

From this expression, one can see that $W_k^\mu(s)$ is the probability of entering A from s during the (residual) time horizon $[N-k, N]$ of length k, under policy μ applied from π.

By (4) and (6), $\mathcal{P}_\pi^\mu(A)$ can be expressed as

$$\mathcal{P}_\pi^\mu(A) = \int_{\mathcal{S}} W_N^\mu(s)\pi(ds), \quad (7)$$

which, for $\pi(ds) = \delta_s(ds)$, reduces to $\mathcal{P}_s^\mu(A) = W_N^\mu(s)$.

Let $\mathcal{F}$ denote the set of functions from $\mathcal{S}$ to $\mathbb{R}$. Define the map $H : \mathcal{S} \times \mathcal{U} \times \Sigma \times \mathcal{F} \to \mathbb{R}$ as follows:

$$H(s, (u, \sigma), Z) := \Big[T_q(q|s, u) \int_{\mathbb{R}^{n(q)}} Z((q, x'))T_x(dx'|s, u) + \sum_{q'\neq q} T_q(q'|s, u) \int_{\mathbb{R}^{n(q')}} Z((q', x'))R(dx'|s, \sigma, q')\Big], \quad (8)$$

$s \in \mathcal{S}$, $(u, \sigma) \in \mathcal{U} \times \Sigma$, and $Z \in \mathcal{F}$.

From definition (5) of W_k^μ, recalling that of T_s in (1), Lemma 1 readily follows.

Lemma 1: Fix a Markov policy $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1})$, $\mu_k : \mathcal{S} \to \mathcal{U} \times \Sigma$, $k = 0, 1, \ldots, N-1$. Then, functions $W_k^\mu : \mathcal{S} \to [0, 1]$, $k = 0, 1 \ldots, N$, can be computed by the recursion:

$$W_{k+1}^\mu(s) = \mathbf{1}_A(s) + \mathbf{1}_{\bar{A}}(s)H(s, \mu_{N-(k+1)}(s), W_k^\mu),\ s \in \mathcal{S},$$

initialized with $W_0^\mu(s) = \mathbf{1}_A(s)$, $s \in \mathcal{S}$. □

B. Multiplicative cost

The probability $\mathcal{P}_\pi^\mu(A)$ defined in (2) can be expressed as

$$\mathcal{P}_\pi^\mu(A) = 1 - p_\pi^\mu(\bar{A}), \quad (9)$$

where $p_\pi^\mu(\bar{A}) := P_\pi^\mu(\mathbf{s}(k) \in \bar{A} \text{ for all } k \in [0, N])$ is the probability of remaining within $\bar{A}$ during the time interval $[0, N]$. 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,

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

From this expression it follows that

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

where the conditional mean under the sign of integral is well defined over the support of the distribution π of $\mathbf{s}(0)$.

Consider a Markov policy $\mu = (\mu_0, \ldots, \mu_{N-1}) \in \mathcal{M}_m$. For each $k \in [0, N], s \in \mathcal{S}$, define $V_k^\mu : \mathcal{S} \to [0, 1]$ as

$$V_k^\mu(s) := \mathbf{1}_{\bar{A}}(s) \int_{\bar{A}^{k+1}} \prod_{h=N-k}^{N-1} T_s^{\mu_h}(ds_{h+1}|s_h)\delta_s(ds_{N-k}),$$

where, for $k = 0$, $\prod_{h=N-k}^{N-1} T_s^{\mu_h}(\cdot|s_h) = 1$, $s_h \in \mathcal{S}$. Note that V_k^μ can be rewritten as

$$\int_{\mathcal{S}^{k+1}} \prod_{l=N-k}^{N} \mathbf{1}_{\bar{A}}(s_l) \prod_{h=N-k}^{N-1} T_s^{\mu_h}(ds_{h+1}|s_h)\delta_s(ds_{N-k}). \quad (12)$$

If $s \in \mathcal{S}$ belongs to the support of the distribution of $\mathbf{s}(N-k)$, then, $E^{\mu}_{\pi}\big[\prod_{l=N-k}^{N} \mathbf{1}_{\bar{A}}(\mathbf{s}(l))|\, \mathbf{s}(N-k) = s\big]$ is well-defined and equal to (12), so that

$$V^{\mu}_{k}(s) = E^{\mu}_{\pi}\big[\prod_{l=N-k}^{N} \mathbf{1}_{\bar{A}}(\mathbf{s}(l))|\, \mathbf{s}(N-k) = s\big] \qquad (13)$$

denotes the probability of remaining inside $\bar{A}$, starting from s, during the (residual) time horizon $[N-k, N]$ of length k, under policy μ applied from π.

By (11) and (13), $p^{\mu}_{\pi}(\bar{A})$ can be expressed as

$$p^{\mu}_{\pi}(\bar{A}) = \int_{\mathcal{S}} V^{\mu}_{N}(s)\pi(ds), \qquad (14)$$

hence $\mathcal{P}^{\mu}_{\pi}(A) = 1 - \int_{\mathcal{S}} V^{\mu}_{N}(s)\pi(ds)$. The following lemma was proven in [3].

Lemma 2: Fix a Markov policy $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1})$, $\mu_k : \mathcal{S} \to \mathcal{U} \times \Sigma$, $k = 0, 1, \ldots, N-1$. Then, functions $V^{\mu}_{k} : \mathcal{S} \to [0,1]$, $k = 0, 1 \ldots, N$, can be computed by the recursion:

$$V^{\mu}_{k+1}(s) = \mathbf{1}_{\bar{A}}(s) H(s, \mu_{N-(k+1)}(s), V^{\mu}_{k}),\ s \in \mathcal{S},$$

initialized with $V^{\mu}_{0}(s) = \mathbf{1}_{\bar{A}}(s)$, $s \in \mathcal{S}$. □

C. Max cost versus multiplicative cost

By $\prod_{h=N-k}^{N} \mathbf{1}_{\bar{A}}(s_h) = 1 - \max_{h \in [N-k, N]} \mathbf{1}_{A}(s_h)$, $k \in [0, N]$, the following can be established.

Lemma 3: Fix a Markov policy $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1})$. For any $k \in [0, N]$, $V^{\mu}_{k}(s) = 1 - W^{\mu}_{k}(s)$, $s \in \mathcal{S}$. □

IV. Probabilistic safe set Computation

For a given policy $\mu \in \mathcal{M}$, different initial conditions s are characterized by a different probability of entering the unsafe set A. If the system starts from an initial condition that corresponds to a probability $\epsilon \in [0,1]$ of entering A, then the system is said to be "safe with probability $1-\epsilon$". We define the *probabilistic safe set* with safety level $1-\epsilon$ associated with policy μ as

$$S^{\mu}(\epsilon) = \{s \in \mathcal{S} : \mathcal{P}^{\mu}_{s}(A) \leq \epsilon\}. \qquad (15)$$

If the policy μ is Markov, then, by (7) and (14) the probabilistic safe set with safety level $1-\epsilon$, $\epsilon \in [0,1]$, defined in (15), can be expressed as the level set of W^{μ}_{N} and V^{μ}_{N}: $S^{\mu}(\epsilon) = \{s \in \mathcal{S} : W^{\mu}_{N}(s) \leq \epsilon\} = \{s \in \mathcal{S} : V^{\mu}_{N}(s) \geq 1-\epsilon\}$.

Suppose that the control policy can be selected so as to minimize the probability of entering A. We can then define the *maximal probabilistic safe set* with safety level $1-\epsilon$:

$$S^{\star}(\epsilon) = \{s \in \mathcal{S} : \inf_{\mu \in \mathcal{M}_m} \mathcal{P}^{\mu}_{s}(A) \leq \epsilon\}. \qquad (16)$$

$S^{\star}(\epsilon)$ is called "maximal" since the safe set $S^{\mu}(\epsilon)$ in (15) associated to a policy $\mu \in \mathcal{M}_m$ satisfies $S^{\mu}(\epsilon) \subseteq S^{\star}(\epsilon)$, for each $\mu \in \mathcal{M}_m$, $\epsilon \in [0,1]$. In this section we show that, for the class of Markov policies, the problem of computing $S^{\star}(\epsilon)$ can be turned into an optimal control problem and solved by dynamic programming.

The calculation of the maximal probabilistic safe set $S^{\star}(\epsilon)$ defined in (16) amounts to finding the infimum over the policies of the probability $\mathcal{P}^{\mu}_{s}(A)$ 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 Markov policy): Let $A \in \mathcal{B}(\mathcal{S})$ be an unsafe set for the DTSHS with state space $\mathcal{S}$. A Markov policy $\mu^* \in \mathcal{M}_m$ is maximally safe if $\mathcal{P}^{\mu^*}_{s}(A) = \inf_{\mu \in \mathcal{M}_m} \mathcal{P}^{\mu}_{s}(A)$, $\forall s \in \bar{A}$. □

In general, a maximally safe policy is not guaranteed to exist. We next provide sufficient conditions for the existence of a maximally safe Markov policy, in terms of both the max cost and the multiplicative cost. The proofs are omitted due to space limitations.

A. Max cost

In the following theorem, we describe an algorithm to compute a maximally safe Markov policy $\mu^* \in \mathcal{M}_m$ based on representation (3) of $\mathcal{P}^{\mu}_{s}(A)$.

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

$$W^{*}_{k+1}(s) = \inf_{(u,\sigma) \in \mathcal{U} \times \Sigma} \big(\mathbf{1}_{A}(s) + \mathbf{1}_{\bar{A}}(s) H(s, (u,\sigma), W^{*}_{k})\big),$$

$s \in \mathcal{S}$, initialized with $W^{*}_{0}(s) = \mathbf{1}_{A}(s)$, $s \in \mathcal{S}$.
Then, $W^{*}_{N}(s) = \inf_{\mu \in \mathcal{M}_m} \mathcal{P}^{\mu}_{s}(A)$, $s \in \mathcal{S}$.

If $\mu^{*}_{k} : \mathcal{S} \to \mathcal{U} \times \Sigma$, $k \in [0, N-1]$, is such that

$$\mu^{*}_{k}(s) = \arg \inf_{(u,\sigma) \in \mathcal{U} \times \Sigma} H(s, (u,\sigma), W^{*}_{N-(k+1)}), \forall s \in \bar{A}, \quad (17)$$

then, $\mu^* = (\mu^{*}_{0}, \ldots, \mu^{*}_{N-1})$ is a maximally safe Markov policy. A sufficient condition for the existence of such a μ^* is that $U_k(s, \lambda) = \{(u,\sigma) \in \mathcal{U} \times \Sigma :\ H(s, (u,\sigma), W^{*}_{k}) \leq \lambda\}$ is compact for all $s \in \bar{A}$, $\lambda \in \mathbb{R}$, $k \in [0, N-1]$. □

B. Multiplicative cost

We now show how to compute a maximally safe Markov policy $\mu^* \in \mathcal{M}_m$, based on equation (9) and on the representation (10) of $p^{\mu}_{s}(\bar{A})$ as a multiplicative cost.

Theorem 2: Define $V^{*}_{k} : \mathcal{S} \to [0,1]$, $k = 0, 1, \ldots, N$, by the recursion:

$$V^{*}_{k+1}(s) = \sup_{(u,\sigma) \in \mathcal{U} \times \Sigma} \mathbf{1}_{\bar{A}}(s) H(s, (u,\sigma), V^{*}_{k}),$$

$s \in \mathcal{S}$, initialized with $V^{*}_{0}(s) = \mathbf{1}_{\bar{A}}(s)$, $s \in \mathcal{S}$.
Then, $V^{*}_{N}(s) = \sup_{\mu \in \mathcal{M}_m} p^{\mu}_{s}(\bar{A})$, $s \in \mathcal{S}$.

If $\mu^{*}_{k} : \mathcal{S} \to \mathcal{U} \times \Sigma$, $k \in [0, N-1]$, is such that

$$\mu^{*}_{k}(s) = \arg \sup_{(u,\sigma) \in \mathcal{U} \times \Sigma} H(s, (u,\sigma), V^{*}_{N-(k+1)}), \forall s \in \bar{A}, \quad (18)$$

then, $\mu^* = (\mu^{*}_{0}, \ldots, \mu^{*}_{N-1})$ is a maximally safe Markov policy. A sufficient condition for the existence of such a μ^* is that $U_k(s, \lambda) = \{(u,\sigma) \in \mathcal{U} \times \Sigma :\ H(s, (u,\sigma), V^{*}_{k}) \geq \lambda\}$ is compact for all $s \in \bar{A}$, $\lambda \in \mathbb{R}$, $k \in [0, N-1]$. □

Remark 1: If $\mathcal{U}$ and Σ are finite sets, then the compactness assumption in Theorems 1 and 2 is not required. □

In view of Theorems 1 and 2, the maximal probabilistic safe set $S^{\star}(\epsilon)$ with safety level $1-\epsilon$ defined in (16) can be determined as either $S^{\star}(\epsilon) = \{s \in \mathcal{S} : W^{*}_{N}(s) \leq \epsilon\}$ or $S^{\star}(\epsilon) = \{s \in \mathcal{S} : V^{*}_{N}(s) \geq 1-\epsilon\}$.

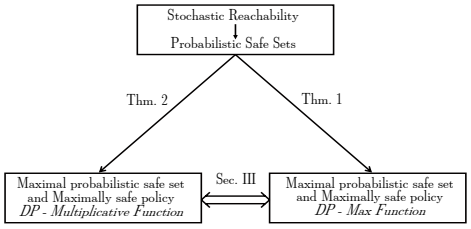

Fig. 1. Problem Interpretation.

Remark 2: The results in this section can be easily extended to the case when the unsafe set A is time varying: $A(k) \in \mathcal{B}(\mathcal{S})$, $k \in [0, N]$. The corresponding recursive expressions should be simply changed by considering at each iteration k the corresponding sets $A(N-k)$ and $\bar{A}(N-k)$, and initializing the recursion with the indicator function of $A(N)$ and $\bar{A}(N)$. □

C. Example

We consider the problem of regulating the temperature of a room during some time horizon $[0, N]$ by a thermostat that can switch a heater on and off.

The DTSHS model $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ is taken from [3]. The discrete component $\mathcal{Q} = \{\texttt{ON}, \texttt{OFF}\}$ represents the heater operating mode. The continuous state component represents the average room temperature, hence $n(q) = 1$, $\forall q \in \mathcal{Q}$. The transition control space is $\mathcal{U} = \{0, 1\}$: "1" means that a switching command is issued to the heater, "0" that no switching command is issued. The reset control space is $\Sigma = \emptyset$.

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. The transition kernel T_x is defined as (see [3] for more details) $T_x(\cdot\,|(q,x),u) = \mathcal{N}(\cdot; m_q(x), \nu^2)$, where $m_{\texttt{OFF}}(x) = x - \frac{a}{C}(x - x_a)\Delta t$, $m_{\texttt{ON}}(x) = m_{\texttt{OFF}}(x) + \frac{r}{C}\Delta t$ and $\nu^2 := \frac{1}{C^2}\Delta t$; a is the average heat loss rate, C is the average thermal capacity of the room, x_a is the constant ambient temperature, r is the rate of heat gain supplied by the heater, and Δt is the time discretization interval. Temperature is measured in Fahrenheit degrees ($^\circ F$) and time in minutes (min). The time horizon has length $N = 600\ min$. The parameters are assigned the following values: $x_a = 10.5^\circ F$, $a/C = 0.1\ min^{-1}$, $r/C = 10^\circ F/min$, $\nu = 0.33^\circ F$, and $\Delta t = 1\ min$.

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

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

where $\alpha \in [0, 1]$ represents the probability of switching from one mode to the other in one time-step. Here, we set $\alpha = 0.8$. The reset map is taken to be equal to T_x: $R(\cdot\,|(q,x),q') = \mathcal{N}(\cdot; m_q(x), \nu^2)$.

We consider the following regulation problem: determine a control law that maximizes the probability that the average room temperature x is driven close to $75^\circ F$ in $t\ min$ starting from any value in the set $(70, 80)^\circ F$, with an admissible excursion of $\pm 1^\circ F$ around $75^\circ F$, and maintained within $75^\circ F \pm 1^\circ F$ thereafter. t is the allowed time to steer the temperature to the desired region and can be specified by the user or chosen by the control designer. We consider the case when $t = 300$. Similar results are obtained for $t = 150$ and $t = 450$. The implementation is done in MATLAB with a discretization step for the temperature equal to $0.05^\circ F$.

The regulation problem can be reformulated as that of computing a maximally safe policy for a time varying "safe" set $\bar{A}(k) = \mathcal{Q} \times \mathcal{X}(k)$, where $\mathcal{X}(k)$ shrinks from the region $(70, 80)^\circ F$ towards the desired region $(74, 76)^\circ F$ during the time interval $[0, 300)\ min$, and then keeps equal to $(74, 76)^\circ F$ in the interval $[300, 600]\ min$.

The results discussed below refer to the following three different evolutions in time of the safe set for the temperature during the time interval $[0, t)$, with $t = 300$: $\mathcal{X}_1(k) = (70, 80)$, $k \in [0, t)$; $\mathcal{X}_2(k) = (70, 80)$, $k \in [0, t/2)$, and $(66 + \frac{8k}{t}, 84 - \frac{8k}{t})$, $k \in [t/2, t)$; $\mathcal{X}_3(k) = \left(70 + \frac{4k}{t}, 80 - \frac{4k}{t}\right)$, $k \in [0, t)$. Correspondingly, $\bar{A}_i(\cdot) = \mathcal{Q} \times \mathcal{X}_i(\cdot)$, $i = 1, 2, 3$.

We determined the maximally safe Markov policies μ_i^*, $i = 1, 2, 3$ and reported them in Figure 2. The plots in the first row refer to $\bar{A}_1(\cdot)$, those in the second row to $\bar{A}_2(\cdot)$, and those in the third row to $\bar{A}_3(\cdot)$. The plots on the left correspond to the OFF mode, and those on the right to the ON mode. Each plot represents the value taken by the binary input u during the time horizon from 0 to 600 min (on the horizontal axis) as a function of the temperature (on the vertical axis). For any time instant $k \in [0, 600]$ only the corresponding safe temperature range is considered. The value 0 ("do not switch") for u is plotted in gray, whereas the value 1 ("switch") is plotted in black. The maximally safe policies are expected to be time-varying during the time interval $[0, 300)$.

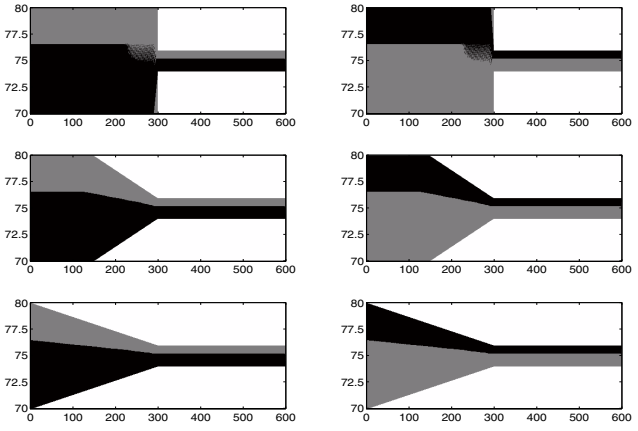

Fig. 2. Maximally safe policy as a function of the temperature and time for the safe sets $\bar{A}_1$, $\bar{A}_2$, and $\bar{A}_3$ (from top to bottom) and $t = 300$. The left (right) column corresponds to the OFF (ON) mode. The darker (lighter) shade indicates that "switch" ("do not switch") is the recommended action.

We computed the probabilities $p_\pi^{\mu_i^*}(\bar{A}_i(\cdot))$, of remaining in the safe sets $\bar{A}_i(\cdot)$, $i = 1, 2, 3$, when the initial distribution

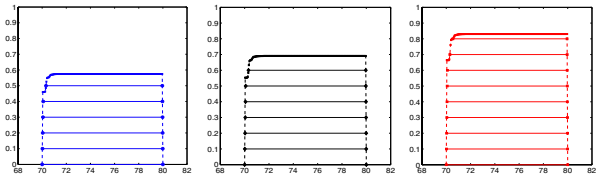

Fig. 3. Maximal probabilistic safe sets for $\mathcal{X}_1(\cdot)$ when $t = 150, 300, 450$ (from left to right) and the heater is initially off. The safety level is reported on the vertical axis, and the temperature value on the horizontal axis.

π is uniform over $\mathcal{Q} \times (70, 80)$. $p_\pi^{\mu_i^*}(\bar{A}_i(\cdot))$ remains almost the same for $\bar{A}_i(\cdot)$, $i = 1, 2, 3$, with the value for $\bar{A}_1(\cdot)$ only marginally higher than the others. This is easily seen since $\bar{A}_3(k) \subseteq \bar{A}_2(k) \subseteq \bar{A}_1(k)$, $k \in [0, 600]$, implies $p_\pi^\mu(\bar{A}_3(\cdot)) \leq p_\pi^\mu(\bar{A}_2(\cdot)) \leq p_\pi^\mu(\bar{A}_1(\cdot))$, for any μ and π.

We furthermore determined the maximal probabilistic safe sets $S_i^*(\epsilon) = \cup_{q \in \{\texttt{OFF},\texttt{ON}\}} \{x \in \mathbb{R} : p_{(q,x)}^{\mu_i^*}(\bar{A}_i(\cdot)) \geq 1 - \epsilon\}$ corresponding to different safety levels $1-\epsilon$, and for different values of the transient length t. In Figure 3 we plotted the subset of $S_i^*(\epsilon)$ corresponding to $q = \texttt{OFF}$ for $i = 1$ and $t = 150, 300, 450$ (the plots when $q = \texttt{ON}$ are similar). Note that, not surprisingly, $S_1^*(\epsilon)$ gets smaller as the required safety level $1 - \epsilon$ grows, and gets larger as t increases. Similar results are obtained for $i = 2$ and $i = 3$.

V. Extensions to the Infinite Horizon Case

We consider a system that is described by a controlled DTSHS model $\mathcal{H}$ (see Definition 1). The sets $\mathcal{M}$ and $\mathcal{M}_m$ of feedback and Markov policies are extension to the infinite horizon case of those introduced in Section II. A Markov policy $\mu \in \mathcal{M}_m$ is said to be stationary if $\mu = (\bar{\mu}, \bar{\mu}, \bar{\mu}, \dots)$, with $\bar{\mu} : \mathcal{S} \to \mathcal{U} \times \Sigma$ universally measurable. The execution of the DTSHS $\mathcal{H}$ associated with some policy μ and initial distribution π is easily obtained by extending Definition 3 to the infinite horizon. Then, the execution $\{\mathbf{s}(k), k \geq 0\}$ associated with $\mu \in \mathcal{M}$ and π is a stochastic process defined on the canonical sample space $\Omega = \mathcal{S}^\infty$, endowed with its product topology $\mathcal{B}(\Omega)$, with probability measure $P_{\pi,\infty}^\mu$ uniquely defined by the transition kernel T_s, the policy μ, and the initial distribution π (see [9, Proposition 7.45]).

For a given policy $\mu \in \mathcal{M}$ and initial distribution π, let

$$\mathcal{P}_{\pi,\infty}^\mu(A) := P_{\pi,\infty}^\mu(\mathbf{s}(k) \in A \text{ for some } k \geq 0),$$

be the probability of entering the unsafe region specified by $A \in \mathcal{B}(\mathcal{S})$. If π is concentrated in a single point s, we use the notation $\mathcal{P}_{s,\infty}^\mu(A)$. The goal is again that of finding an optimal Markov policy that singles out the *maximal probabilistic safe set* $S_\infty^\star(\epsilon) := \{s \in \mathcal{S} : \inf_{\mu \in \mathcal{M}_m} \mathcal{P}_{\pi,\infty}^\mu(A) \leq \epsilon\}$ with safety level $1 - \epsilon$.

We aim at computing this maximally safe policy by means of a dynamic programming scheme. In addition, as it is reasonable in an infinite horizon setting for a time-invariant system, we try to investigate if such a policy can be selected among the stationary Markov policies.

In the following, we shall focus on the interpretation based on the expression for the probability $\mathcal{P}_{\pi,\infty}^\mu(A)$ in terms of the max cost: $\mathcal{P}_{\pi,\infty}^\mu(A) = E_{\pi,\infty}^\mu\big[\max_{k \geq 0} \mathbf{1}_A(\mathbf{s}(k))\big]$.

Unlike the additive-cost, the max cost framework is not of wide usage. Extending the results developed for the infinite horizon additive cost case to the infinite horizon max cost case requires some attention regarding the following aspects: it is first necessary to take care of the asymptotic properties of the max cost function at the limit; and the measurability properties of the limit function, its behavior when minimized, the existence and properties of the optimal argument have to be carefully assessed.

An iterative procedure to compute $\mathcal{P}_{\pi,\infty}^\mu(A)$ is again possible. Conditions that yield a stationary optimal Markov policy can also be provided. The proof of this result is omitted due to space limitations.

Theorem 3: Define the maps $W_k^* : \mathcal{S} \to [0, 1]$, $k \geq 0$, by the recursion:

$$W_{k+1}^*(s) = \inf_{(u,\sigma) \in \mathcal{U} \times \Sigma} \big(\mathbf{1}_A(s) + \mathbf{1}_{\bar{A}}(s) H(s, (u, \sigma), W_k^*)\big),$$

$s \in \mathcal{S}$, initialized with $W_0^*(s) = \mathbf{1}_A(s)$, $s \in \mathcal{S}$.
Suppose that $\exists \bar{k} \geq 0$ s.t. $\{(u, \sigma) \in \mathcal{U} \times \Sigma : H(s, (u, \sigma), W_k^*) \leq \lambda\}$ is compact for all $s \in \bar{A}$, $\lambda \in \mathbb{R}$, $k \geq \bar{k}$. Then, $W_\infty^*(s) = \inf_{\mu \in \mathcal{M}_m} \mathcal{P}_{s,\infty}^\mu(A)$, $s \in \mathcal{S}$. Furthermore, there exists a maximally safe stationary Markov policy $\mu^* = (\bar{\mu}^*, \bar{\mu}^*, \dots)$, with $\bar{\mu}^* : \mathcal{S} \to \mathcal{U} \times \Sigma$, given by

$$\bar{\mu}^*(s) = \arg \inf_{(u,\sigma) \in \mathcal{U} \times \Sigma} H(s, (u, \sigma), W_\infty^*), \forall s \in \bar{A}. \qquad \square$$

VI. Future Work

There appear to be very promising directions for this research, from both a theoretical and an application-oriented viewpoint. Regarding the latter aspect, the idea of embedding further performance specifications in the control design procedure, while guaranteeing a reachability specification, appears an interesting possibility to investigate.

References

[1] D. P. Bertsekas and I. B. Rhodes, "On the minimax reachability of target sets and target tubes," *Automatica*, vol. 7, pp. 233–247, 1971.

[2] J. Hu, M. Prandini, and S. Sastry, "Aircraft conflict prediction in the presence of a spatially correlated wind field," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, pp. 326–340, 2005.

[3] S. Amin, A. Abate, M. Prandini, J. Lygeros, and S. Sastry, "Reachability analysis for controlled discrete time stochastic hybrid systems," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science 3927, J. Hespanha and A. Tiwari, Eds. Springer Verlag, 2006, pp. 49–63.

[4] C. Tomlin, J. Lygeros, and S. Sastry, "Synthesizing controllers for nonlinear hybrid systems," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science 1386, T. Henzinger and S. Sastry, Eds. Springer Verlag, 1998, pp. 360–373.

[5] A. Balluchi, L. Benvenuti, M. D. Di Benedetto, G. M. Miconi, U. Pozzi, T. Villa, H. Wong-Toi, and A. L. Sangiovanni-Vincentelli, "Maximal safe set computation for idle speed control of an automotive engine," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science 1790, N. Lynch and B. Krogh, Eds. Springer Verlag, 2000, pp. 32–44.

[6] D. P. Bertsekas, "Infinite-time reachability of state-space regions using feedback control," *IEEE Transactions on Automatic Control*, vol. AC-17, no. 5, pp. 604–613, 1972.

[7] B. Picasso and A. Bicchi, "Control synthesis for practical stabilization of quantized linear systems," *Ren. Sem. Mat. Universita' di Torino*, vol. 63, no. 4, pp. 397–410, 2005.

[8] M. Puterman, *Markov decision processes*. John Wiley & Sons, Inc, 1994.

[9] D. P. Bertsekas and S. E. Shreve, *Stochastic optimal control: the discrete-time case*. Athena Scientific, 1996.