Automatica 44 (2008) 2724–2734

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems☆

Alessandro Abate [a,*], Maria Prandini [b], John Lygeros [c], Shankar Sastry [a]

[a] Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, USA

[b] Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy

[c] Automatic Control Laboratory, ETH Zurich, Switzerland

ARTICLE INFO

Article history:
Received 13 August 2007
Received in revised form
18 December 2007
Accepted 22 March 2008
Available online 9 October 2008

Keywords:
Hybrid systems
Stochastic systems
Reachability
Safety
Optimal control
Dynamic programming

ABSTRACT

In this work, probabilistic reachability over a finite horizon is investigated for a class of discrete time stochastic hybrid systems with control inputs. A suitable embedding of the reachability problem in a stochastic control framework reveals that it is amenable to two complementary interpretations, leading to dual algorithms for reachability computations. In particular, the set of initial conditions providing a certain probabilistic guarantee that the system will keep evolving within a desired 'safe' region of the state space is characterized in terms of a value function, and 'maximally safe' Markov policies are determined via dynamic programming. These results are of interest not only for safety analysis and design, but also for solving those regulation and stabilization problems that can be reinterpreted as safety problems. The temperature regulation problem presented in the paper as a case study is one such case.

1. Introduction

Engineering systems like air traffic management systems and infrastructure networks, and natural systems like biological networks exhibit complex behaviors arising from the interactions between heterogeneous components, and are intrinsically characterized by hybrid dynamics. The uncertainty affecting the interleaved discrete and continuous evolutions results in a different likelihood of the system trajectories, thus naturally leading to stochastic hybrid models. The analysis and control of a stochastic hybrid system is generally more challenging than that of a deterministic hybrid system. In this paper, we concentrate on the reachability problem for discrete time stochastic hybrid systems.

Reachability is an important topic in classical control theory. In general terms, a reachability problem consists of evaluating whether the state of a system will reach a certain set during some time horizon, starting from a given set of initial conditions and possibly subject to a control input. This problem is of interest, for instance, in those safety problems where the system should be kept outside an unsafe region of the state space and the control input can be chosen so as to avoid this unsafe region. In a deterministic setting, reachability is a yes/no problem, where one evaluates if starting from a given set of initial states the system will reach a certain set or not. In a stochastic setting, the different trajectories originating from each initial state have a different likelihood and one can then evaluate what is the probability that the system will reach the assigned set starting from a certain initial distribution over the set of initial states. In safety problems where the evolution of the system can be influenced by some control input, one should select it appropriately so as to minimize the probability that the state of the system will enter the unsafe set.

Much investigation has been done on reachability analysis for system verification, where the problem consists in verifying if some designed system satisfies certain reachability specifications encoding a correct/desired behavior.

In the case of deterministic systems, 'model checking' is the most commonly adopted technique for system verification, where reachability specifications are verified by constructing reachable sets based on a model of the system. In the hybrid systems

☆ 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 'Identification and Adaptive Control of Industrial Systems', and by the NSF grant CCR-0225610. This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor George Yin under the direction of Editor Ian Richard Petersen.

* Corresponding author.
E-mail addresses: aabate@eecs.berkeley.edu (A. Abate), prandini@elet.polimi.it (M. Prandini), lygeros@control.ee.ethz.ch (J. Lygeros), sastry@eecs.berkeley.edu (S. Sastry).

0005-1098/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2008.03.027

case, set representation and propagation by continuous flow is generally difficult, and termination of the algorithm for reachable set computation is not guaranteed since the state space is not finite (Tomlin, Mitchell, Bayen, & Oishi, 2003). Decidability results have been proven to hold only for certain classes of hybrid systems by using discrete abstraction consisting in building a finite automaton that is 'equivalent' (bisimilar) to the original hybrid system for the purpose of verification (Alur, Henzinger, Lafferriere, & Pappas, 2000).

In the case of complex dynamics, some approximation methods are needed for reachability computations. Two main approaches have been introduced to this purpose: seeking for an abstraction of the system that would yield a simpler model for solving the original reachability problem, or adopting an approximation of sets that can be easily represented and propagated through the system dynamics. In the first approach, an approximate simulation relation is introduced to obtain an abstraction of the original system (Girard, Julius, & Pappas, 2006). In the second approach, over-approximations by ellipsoids (Kurzhanski & Varaiya, 2002), polyhedra (Asarin, Bournez, Dang, & Maler, 2000), zonotopes (Girard, 2005), oriented rectangular polytopes (Stursberg & Krogh, 2003) were proposed, or, alternatively, asymptotic approximations of reachable sets that converge to the true reachable sets as some accuracy parameter tends to zero. Level set methods (Mitchell & Tomlin, 2000) and gridding (Belta et al., 2004) techniques belong to this latter category.

A connection of reachability (and related concepts, such as safety or viability) with optimal control for deterministic problems has been pointed out in Lygeros (2004). The connection between rechability, safety and dynamic games for deterministic hybrid systems has been stressed in Lygeros, Tomlin, and Sastry (1999) and Tomlin, Lygeros, and Sastry (1998), where it is mostly applied to air traffic management problems.

Reachability for stochastic hybrid systems has been a very recent focus of research. Most of the approaches consider the problem of reachability analysis for continuous time stochastic hybrid systems (CTSHS) without a control input. The theory of CTSHS, progressively developed since the early contributions in Davis (1993), Ghosh, Araposthasis, and Marcus (1997) and Hu, Lygeros, and Sastry (2000), is used in Bujorianu (2004) to address theoretical issues regarding the measurability of the reachability events. In this reference, the theory of Dirichlet forms associated with a right-Markov process is employed in studying a probabilistic reachability problem, and upper bounds for the reach set probabilities are derived. The contributions in Hu, Prandini, and Sastry (2005) and Prandini and Hu (2006a,b) address the reachability problem using a Markov chain approximation, (Kushner & Dupuis, 2001), to compute the probability of entering some assigned set, and apply the concept to air traffic control studies. Probabilities rather than sets are propagated through the approximating Markov chain transition kernel. In the same spirit, model checkers for verifying probabilistic reachability specifications of Markov chains have been developed, (Katoen, 2006). From a different perspective, in Prajna, Jadbabaie, and Pappas (2007) certain functions of the state of the system known as barrier certificates are used to compute an upper bound on the probability of reaching a set. The approach in Digailova and Kurzhanski (2005) is unique in introducing a 'mean-square' definition for the concept of reachability. In Mitchell and Templeton (2005) the control case is considered in a rather general game theoretical framework, and a reachability problem is introduced as the solution of a Hamilton-Jacobi-Isaacs partial differential equation. Lygeros and Watkins (2003) and Prandini, Hu, Lygeros, and Sastry (2000) compute the reachability probability using randomized algorithms, motivated by air traffic control applications.

In this study, we adopt a discrete time point of view in order to gain a deeper understanding of the theoretical and computational aspects associated with reachability and safety problems for stochastic hybrid systems, while avoiding technical measurability issues that arise in the continuous time case. In particular, we develop a methodology to compute and maximize the probability of maintaining the state of the system within a certain 'safe' region for a class of discrete time stochastic hybrid system (DTSHS) whose dynamics can be influenced by a control input. Unlike previous methods, the safe set can be time-varying, which allows us to generalize the approach towards problems of regulation and stabilization, (Bertsekas, 1972; Picasso & Bicchi, 2005), by appropriately reinterpreting them as safety problems.

The proposed methodology is based on formulating the reachability problem as a stochastic optimal control problem. Based on the expression of the probability that the state of the controlled system evolves within the safe region as a multiplicative cost, dynamic programming (DP) can be used to compute the Markov policy maximizing this cost, and also the 'maximally safe' sets corresponding to different safety levels. These sets characterize the initial conditions for the system, such that there exists a Markov policy capable of maintaining the state of the system within the safe set with a probability greater than a prescribed safety level (see Balluchi et al. (2000) and Tomlin et al. (1998) for a similar notion in the deterministic case).

Adopting a dual perspective, where the objective is that of minimizing the probability that the system will exit the safe set, we again formulate the reachability problem as a stochastic optimal control problem, but this time with a cost that is the maximum of a function of the state over the time horizon. DP is shown to be still effective in this case to determine probabilistic maximally safe sets for Markov policies. In fact, the value functions for the multiplicative cost and the max cost can be properly put in relation, thus formalizing the intuition that the two viewpoints for reachability are complementary to each other.

It is perhaps worth noting that, once the reachability problem is formulated within the stochastic optimal control framework, one could –in principle– refer to the related literature for its solution. The key reference for optimal stochastic control of discrete time stochastic systems is Bertsekas and Shreve (1996), adopting the dynamic programming approach. However, the results in Bertsekas and Shreve (1996) cannot be directly applied to the case under study but need to be extended to jointly take care of the fact that the state is hybrid and characterized by complex interacting continuous and discrete dynamics, and that the objective function is a multiplicative/max cost. These two coexisting aspects require a thorough analysis involving non trivial measurability considerations and a rethinking of the inspiring proof techniques in Bertsekas and Shreve (1996).

Throughout the paper, we shall use a room temperature regulation problem as running example to illustrate the DTSHS model formalism and the approach to reachability. This case study is inspired by one of the benchmark problems proposed in Fehnker and Ivančić (2004).

The paper unfolds as follows: Section 2 introduces our reference DTSHS model. This model is inspired by other stochastic hybrid systems models previously introduced in Bujorianu and Lygeros (2004), Davis (1993), Ghosh et al. (1997) and Hu et al. (2000) for the continuous time case. In Section 3, the temperature regulation problem is outlined and the controlled system is modeled as a DTSHS. In Section 4, the notion of stochastic reachability for a DTSHS is introduced and the problem of determining probabilistic maximally safe sets and maximally safe Markov policies is addressed according to the two complementary viewpoints, leading to DP schemes using respectively a multiplicative and a max cost function. The proposed methodology is applied to the temperature regulation problem in Section 5. Finally, extensions of the above study are discussed and concluding remarks are drawn.

2. Stochastic hybrid system model

We define a DTSHS as the discrete time counterpart of the general continuous time model described in Bujorianu and Lygeros (2004), extending in expressiveness previous continuous time models (Davis, 1993; Ghosh et al., 1997; Hu et al., 2000).

The state of a DTSHS is characterized by a discrete and a continuous component. The discrete state component takes on values in a countable set of modes $\mathcal{Q}$. The continuous state space in each mode $q \in \mathcal{Q}$ is given by the Euclidean space $\mathbb{R}^{n(q)}$, whose dimension $n(q)$ 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 $\cup_{q\in\mathcal{Q}}\{q\} \times A_q$, where A_q is a Borel set in $\mathbb{R}^{n(q)}$. The fact that $\mathcal{S}$ can be endowed with a metric that is equivalent to the usual Euclidean metric when restricted to each domain $\mathbb{R}^{n(q)}$ (Davis, 1993), shows that $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$ is a Borel space, i.e. homeomorphic to a Borel subset of a complete separable metric space.

The continuous state of a DTSHS evolves according to a probabilistic law that depends on the actual operating mode. A discrete transition from the current operating mode to a different one may occur during the continuous state evolution, again according to some probabilistic law. This will in turn cause a modification of the probabilistic law governing the continuous state dynamics. A control input can affect the discrete and continuous evolution of the system. Moreover, after a discrete transition has occurred, the continuous state component is subject to a probabilistic reset that is also influenced by some control input. We distinguish this latter input from the former one, naming them respectively *reset* and *transition input*.

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

To describe the semantics of a DTSHS, we need to specify an initial condition $s_0 \in \mathcal{S}$ (which may be sampled from an initial probability distribution) and how the reset and transition inputs are chosen. Here, we consider a DTSHS evolving over a finite time horizon $[0, N]$, with inputs chosen according to a Markov policy.

Definition 2 (*Markov Policy*). A Markov policy π for a DTSHS $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ is a sequence $\pi = (\mu_0, \mu_1, \mu_2, \ldots)$ of universally measurable maps $\mu_k : \mathcal{S} \to \mathcal{U} \times \Sigma$, $k = 0, 1, 2, \ldots$, from the hybrid state space $\mathcal{S} = \cup_{q\in\mathcal{Q}}\{q\} \times \mathbb{R}^{n(q)}$ to the control input space $\mathcal{U} \times \Sigma$. □

We denote the set of Markov policies as $\mathcal{M}_m$.

We recall that a function $\mu_k : \mathcal{S} \to \mathcal{U} \times \Sigma$ is universally measurable if the inverse image of every Borel set is measurable with respect to every complete probability measure on $\mathcal{S}$ that measures all Borel subsets of $\mathcal{S}$. This measurability condition is weaker than the Borel measurability condition and is needed to assess properties which hold uniformly in the initial condition s_0, (Bertsekas & Shreve, 1996).

The semantics of a DTSHS can be algorithmically defined through the notion of *execution*. In the sequel, we shall use boldface to denote random variables and normal typeset to denote sample values.

Definition 3 (*Execution*). Consider a DTSHS $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$ and a time horizon $[0, N]$. 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 Markov policy $\pi \in \mathcal{M}_m$ and an initial condition $s_0 = (q_0, x_0) \in \mathcal{S}$ if its sample paths are obtained according to the DTSHS algorithm:

set $\mathbf{q}(0) = q_0$, $\mathbf{x}(0) = x_0$, and $k = 0$;
while $k < N$ do
  set $(u_k, \sigma_k) = \mu_k((q_k, x_k))$;
  extract from $\mathcal{Q}$ a value q_{k+1} for $\mathbf{q}(k + 1)$ 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 $R(\cdot\,|(q_k, x_k), \sigma_k, q_{k+1})$;
  $k \to k + 1$;
end. □

By appropriately defining the discrete transition kernel T_q, it is possible to model the *spontaneous jumps* that *may* occur during the continuous state evolution, as well as the *forced jumps* that *must* occur when the continuous state exits some prescribed domain.

As for the spontaneous transitions, if a discrete transition from q to $q' \neq q$ is enabled at $(q, x) \in \mathcal{S}$ by the control input $u \in \mathcal{U}$, then this can be encoded by the condition $T_q(q'|(q, x), u) > 0$.

As for the forced transitions, the *invariant set* $Inv(q)$ associated with mode $q \in \mathcal{Q}$, namely the set of all the admissible values for the continuous state within q, can be expressed in terms of T_q by forcing $T_q(q|(q, x), u)$ to be equal to zero for all the continuous state values $x \in \mathbb{R}^{n(q)}$ outside $Inv(q)$, irrespectively of the value of the control input $u \in \mathcal{U}$. Thus $Inv(q) := \mathbb{R}^{n(q)} \setminus \{x \in \mathbb{R}^{n(q)} : T_q(q|(q, x), u) = 0, \forall u \in \mathcal{U}\}$, and as soon as $x \notin Inv(q)$ while the system evolves in mode q, a jump from q to some $q' \neq q$ is forced. Then, unlike the continuous time model in Bujorianu and Lygeros (2004), spatial guards here are implicitly defined through T_q.

Introduce 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 $\mathcal{S} \times \mathcal{U} \times \Sigma \times \mathcal{Q}$:

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

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 $(\mathbb{R}^{n(q')}, \mathcal{B}(\mathbb{R}^{n(q')}))$. The kernel τ_x is used in the DTSHS algorithm to randomly select a value for the continuous state at time $k+1$, given the values taken by the hybrid state and the control input at time k, and that of the discrete state at time $k + 1$.

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 $\mathcal{S} \times \mathcal{U} \times \Sigma$:

$$T_s((\cdot, q)\,|s, (u, \sigma)) = \tau_x(\cdot\,|s, u, \sigma, q)T_q(q|s, u), \quad q \in \mathcal{Q}, \tag{1}$$

which assigns to each $s \in \mathcal{S}, (u, \sigma) \in \mathcal{U} \times \Sigma$ a probability measure on the Borel space $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$.

Then, the DTSHS algorithm in Definition 3 can be rewritten in a more compact form as:

DTSHS algorithm

set $\mathbf{s}(0) = s_0$ *and* $k = 0$;

while $k < N$ *do*

set $(u_k, \sigma_k) = \mu_k(s_k)$;

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) (Puterman, 1994). This will be referred to in the following as *embedded controlled Markov process*.

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 $s_0 \in \mathcal{S}$ and $\pi \in \mathcal{M}_m$ 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 $P^{\pi}_{s_0}$ uniquely defined by the transition kernel T_s, the policy $\pi \in \mathcal{M}_m$, and the initial condition $s_0 \in \mathcal{S}$ (Bertsekas & Shreve, 1996, Proposition 7.45). From the embedded Markov process representation of a DTSHS it also follows that the execution of a DTSHS associated with a Markov policy $\pi = (\mu_0, \mu_1, \ldots) \in \mathcal{M}_m$ and an initial condition s_0 is an inhomogeneous Markov process with one-step transition kernels $T_s(\cdot|s, \mu_k(s))$, $k = 0, 1, \ldots, N-1$. In the sequel we shall use the more compact notation $T_s^{\mu_k}(\cdot|s)$ for $T_s(\cdot|s, \mu_k(s))$.

3. Case study: Temperature regulation — model

We consider the problem of regulating the temperature in r rooms. We suppose that each room can be warmed by a single heater and that at most one heater can be active at a time. The problem consists in designing a control switching strategy that decides which room should be heated, based on the measurements of the temperatures of r rooms, so as to maintain the temperature of each room within a prescribed interval.

We next describe the controlled system through a DTSHS model $\mathcal{H} = (\mathcal{Q}, n, \mathcal{U}, \Sigma, T_x, T_q, R)$, whereas the precise formulation of the temperature control synthesis problem is postponed to Section 5.

The system configuration is naturally described by a hybrid state, whose discrete component represents which of the r rooms is being heated, and whose continuous component represents the average temperature in each of the r rooms. The discrete state space can then be defined as $\mathcal{Q} = \{\mathtt{ON}_1, \mathtt{ON}_2, \ldots, \mathtt{ON}_r, \mathtt{OFF}\}$, where in mode $\mathtt{ON}_i$ it is room i to be heated and in mode OFF no room is heated. The map $n : \mathcal{Q} \to \mathbb{N}$, defining the dimension of the continuous component of the hybrid state space, is the constant map $n(q) = r, \forall q \in \mathcal{Q}$.

A transition input dictates which room is to be heated, whereas no reset input affects the value of the temperature after a discrete transition occurs–intuitively, the temperature of a room is instantaneously unaltered after a discrete switch. Thus, the reset control space is trivial, $\Sigma = \{0\}$, and the transition control space is $\mathcal{U} = \{\mathtt{SW}_1, \mathtt{SW}_2, \ldots, \mathtt{SW}_r, \mathtt{SW}_{\mathtt{OFF}}\}$, where $\mathtt{SW}_i$ and $\mathtt{SW}_{\mathtt{OFF}}$ correspond to the command of heating room i and heating no room, respectively. Note that if the system is operating in mode $\mathtt{ON}_i$ and the transition control input $\mathtt{SW}_i$ is applied, it means that the heater in room i should stay active, thus leaving the current situation unchanged.

Regarding the dynamics of the continuous state $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r)$, we model the evolution of the average temperature $\mathbf{x}_i$ in room i by the following linear stochastic difference equation:

$$\mathbf{x}_i(k+1) = \mathbf{x}_i(k) + \sum_{j\neq i} a_{ij}\left(\mathbf{x}_j(k) - \mathbf{x}_i(k)\right) + b_i\left(x_a - \mathbf{x}_i(k)\right) + c_i h_i(k) + \mathbf{n}_i(k), \tag{2}$$

which is obtained by discretizing, via the constant-step Euler-Maruyama scheme with discretization step Δ, the continuous time equations in Malhame and Chong (1985).

The meaning of the terms appearing in Eq. (2) is the following. x_a is the ambient temperature, which is assumed to be constant. b_i, a_{ij}, and c_i are non negative constants representing the average heat loss rates of room i to the ambient (b_i) and to room $j \neq i$ (a_{ij}), and the rate of heat supplied by the heater in room i (c_i), all normalized with respect to the average thermal capacity of room i and rescaled by Δ. The values taken by the a_{ij} constants reflect the rooms layout, for instance $a_{ij} = 0$ if rooms i and j are not adjacent. The term $h_i(k)$ is a Boolean function equal to 1 if $\mathbf{q}(k) = \mathtt{ON}_i$ (i.e. if it is room i to be heated at time k), and equal to 0 otherwise. Furthermore, the disturbance $\{\mathbf{n}_i(k), k = 0, \ldots, N-1\}$ affecting the temperature evolution is a sequence of i.i.d Gaussian random variables with zero mean and variance ν^2 proportional to Δ. We suppose for simplicity that the disturbances affecting the temperatures of different rooms are independent.

Let $\mathcal{N}(\cdot; m, V)$ denote the probability measure over $(\mathbb{R}^r, \mathcal{B}(\mathbb{R}^r))$ associated with a Gaussian density function with mean m and covariance matrix V. Then, the continuous transition kernel T_x (implicitly defined via the stochastic difference equation (2)) can be expressed as follows:

$$T_x(\cdot\,|(q, x), u) = \mathcal{N}(\cdot; x + \Xi x + \Gamma(q), \nu^2 I), \tag{3}$$

where Ξ is a square matrix of size r, $\Gamma(q)$ is an r-dimensional column vector that depends on $q \in \mathcal{Q}$, and I is the identity matrix of size r.

The element in row i and column j of matrix Ξ is given by $[\Xi]_{ij} = a_{ij}$, if $j \neq i$, and $[\Xi]_{ij} = -b_i - \sum_{k\neq i, k\in\mathcal{Q}} a_{ik}$, if $j = i$, for any $i = 1, \ldots, r$. As for the vector $\Gamma(q)$, its ith component is $[\Gamma(q)]_i = b_i x_a + c_i$, if $q = \mathtt{ON}_i$, and $[\Gamma(q)]_i = b_i x_a$, if $q \in \mathcal{Q} \setminus \{\mathtt{ON}_i\}$, for any $i = 1, \ldots, r$.

The independence of the sequence of r-dimensional random variables $\{(\mathbf{n}_1(k), \mathbf{n}_2(k), \ldots, \mathbf{n}_r(k)), k = 0, 1, \ldots, N-1\}$ is required for the temperature at time $k+1$ to be conditionally independent on the past, given the hybrid state and the transition input at time k, which allows the description of the temperature evolution within each mode by a transition kernel T_x. Instead, the assumption that the disturbances $\mathbf{n}_i$, $i = 1, 2, \ldots, r$, affecting the temperatures of different rooms are independent can be easily removed. If this were not the case, then, the only modification to be introduced is that the covariance matrix appearing in (3), representing the covariance of the r-dimensional Gaussian random variable $(\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_r)$, would not be diagonal.

Note that the transition kernel T_x governing the temperature evolution within a mode does not depend on the value u taken by the transition control input. This follows from the fact that the transition control input does not affect the temperature dynamics described in (2). We shall then use the notation $T_x(\cdot|(q, x))$ in place of $T_x(\cdot|(q, x), u)$.

We assume that during the time step when a discrete transition occurs, say from mode $\mathtt{ON}_i$ to mode $\mathtt{ON}_j$, the temperature keeps evolving according to the dynamics characterizing the starting condition $\mathtt{ON}_i$. This is modeled by defining the reset kernel as $R(\cdot\,|(q, x), q') = T_x(\cdot\,|(q, x))$, $q, q' \in \mathcal{Q}$, $x \in \mathbb{R}^r$. Note that R does not depend on the value σ of the reset control input, since in this example the reset control space is empty.

The transition control input affects the discrete state evolution through the discrete transition kernel T_q. In this example, discrete

transitions are not influenced by the value taken by the continuous state component, so that we can take $T_q : \mathcal{Q} \times \mathcal{Q} \times \mathcal{U} \to [0, 1]$, with $T_q(q'|q, u)$ representing the probability that mode q' is the successor of mode q when the transition control input u is applied. For ease of notation we set

$$T_q(q'|q, u) = \alpha_{qq'}(u), \quad q, q' \in \mathcal{Q}. \tag{4}$$

Thus, the discrete state evolves according to a (finite state and finite input) controlled Markov chain, with controlled transition probability from state q to state q' under input u given by $\alpha_{qq'}(u)$ in (4).

Not all the transitions may actually occur from a node q. For instance, if the control input value $u = \mathtt{SW}_i$ is applied at node $q = \mathtt{ON}_i$, then with probability one the successor node is $q' = \mathtt{ON}_i$, because room i is currently heated and the command of heating room i is issued. If everything worked perfectly, then, the control input $u = \mathtt{SW}_i$ would lead to node $q' = \mathtt{ON}_i$ from any node q, and, similarly, $u = \mathtt{SW}_{\mathtt{OFF}}$ would lead to $q' = \mathtt{OFF}$ from any node q. The definition of the controlled transition probabilities $\{\alpha_{qq'}(u),\ q, q' \in \mathcal{Q}\}$ associated with the different $u \in \mathcal{U}$ offers the possibility of encoding delays or faulty behaviors in the commutations as well as structural constraints imposing, for instance, that the heat can be conveyed only from a room to a contiguous one.

4. Probabilistic reachability and safety

We consider the problem of determining the probability that the state of a DTSHS $\mathcal{H}$ will remain within a certain 'safe' set during a time horizon $[0, N]$ starting from s_0, under some control policy $\pi \in \mathcal{M}_m$. This probabilistic safety problem can be clearly classified as a stochastic reachability analysis problem.

Let the Borel set $A \in \mathcal{B}(S)$ represent a safe set. Our goal is setting up a reachability computation procedure to determine the probability that the execution associated with the Markov policy $\pi \in \mathcal{M}_m$ and the initial condition s_0 will remain within A during the time horizon $[0, N]$:

$$\mathrm{p}^{\pi}_{s_0}(A) := P^{\pi}_{s_0}\{\mathbf{s}(k) \in A \text{ for all } k \in [0, N]\}. \tag{5}$$

If $\mathrm{p}^{\pi}_{s_0}(A) \geq 1 - \epsilon, \epsilon \in [0, 1]$, we say that the system initialized at s_0 is safe with at least probability $1 - \epsilon$ under policy π.

Different initial conditions are characterized by different values of the probability $\mathrm{p}^{\pi}_{s_0}(A)$. Fix $\epsilon \in [0, 1]$. We define as *probabilistic safe set* with safety level $1 - \epsilon$ under policy π the set

$$S^{\pi}(\epsilon) = \{s_0 \in \mathcal{S} : \mathrm{p}^{\pi}_{s_0}(A) \geq 1 - \epsilon\} \tag{6}$$

of those initial conditions s_0 that correspond to a probability $\mathrm{p}^{\pi}_{s_0}(A)$ of remaining within the safe set A that is greater than or equal to $1 - \epsilon$.

If for any initial condition $s_0 \in \mathcal{S}$ the control policy $\pi \in \mathcal{M}_m$ can be selected so as to maximize the probability of staying within A, then, we can define the set

$$S^{\star}(\epsilon) = \{s_0 \in \mathcal{S} : \sup_{\pi \in \mathcal{M}_m} \mathrm{p}^{\pi}_{s_0}(A) \geq 1 - \epsilon\}. \tag{7}$$

By comparing the expressions for $S^{\pi}(\epsilon)$ and $S^{\star}(\epsilon)$, it is easily seen that $S^{\pi}(\epsilon) \subseteq S^{\star}(\epsilon)$, for each $\pi \in \mathcal{M}_m$ and for any $\epsilon \in [0, 1]$, since in fact we are exploiting the best available control to achieve the ϵ-dependent reachability specification for the largest possible subset of the hybrid state space. The set $S^{\star}(\epsilon)$ is named the *maximal probabilistic safe set* with safety level $1 - \epsilon$. Computing $S^{\star}(\epsilon)$ involves solving an optimization problem, and is a more challenging goal than computing $\mathrm{p}^{\pi}_{s_0}(A)$ and $S^{\pi}(\epsilon)$.

Note that the probability $\mathrm{p}^{\pi}_{s_0}(A)$ defined in (5) can be expressed as

$$\mathrm{p}^{\pi}_{s_0}(A) = 1 - \mathrm{P}^{\pi}_{s_0}(\bar{A}), \tag{8}$$

where $\bar{A}$ is the complement of A in $\mathcal{S}$, and

$$\mathrm{P}^{\pi}_{s_0}(\bar{A}) := P^{\pi}_{s_0}\{\mathbf{s}(k) \in \bar{A} \text{ for some } k \in [0, N]\} \tag{9}$$

is the probability of entering $\bar{A}$ during the time interval $[0, N]$. This leads to the following alternative expressions for $S^{\pi}(\epsilon)$ and $S^{\star}(\epsilon)$:

$$S^{\pi}(\epsilon) = \{s_0 \in \mathcal{S} : \mathrm{P}^{\pi}_{s_0}(\bar{A}) \leq \epsilon\} \tag{10}$$

$$S^{\star}(\epsilon) = \{s_0 \in \mathcal{S} : \inf_{\pi \in \mathcal{M}_m} \mathrm{P}^{\pi}_{s_0}(\bar{A}) \leq \epsilon\}. \tag{11}$$

In the rest of the section, we show that (i) the problem of computing $\mathrm{p}^{\pi}_{s_0}(A)$, $\mathrm{P}^{\pi}_{s_0}(\bar{A})$, and $S^{\pi}(\epsilon)$ for $\pi \in \mathcal{M}_m$ can be solved by using a backward iterative procedure; and that (ii) the problem of computing $S^{\star}(\epsilon)$ can be reduced to that of solving an optimal control problem by the application of dynamic programming.

These results are obtained by representing $\mathrm{p}^{\pi}_{s_0}(A)$ as a multiplicative cost function, and $\mathrm{P}^{\pi}_{s_0}(\bar{A})$ as a max cost function.

Let $\mathbf{1}_C : \mathcal{S} \to \{0, 1\}$ denote the indicator function of set $C \subseteq \mathcal{S}$: $\mathbf{1}_C(s) = 1$, if $s \in C$, and $= 0$, if $s \notin C$.

Multiplicative Cost. Observe that

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

where $s_k \in \mathcal{S}$, $k \in [0, N]$. Then, $\mathrm{p}^{\pi}_{s_0}(A)$ in (5) can be expressed as the expectation with respect to the probability measure $P^{\pi}_{s_0}$ of the Bernoulli random variable $\prod_{k=0}^{N} \mathbf{1}_A(\mathbf{s}(k))$:

$$\mathrm{p}^{\pi}_{s_0}(A) = E^{\pi}_{s_0}\left[\prod_{k=0}^{N} \mathbf{1}_A(\mathbf{s}(k))\right]. \tag{12}$$

Max Cost. Since

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

where $s_k \in \mathcal{S}$, $k \in [0, N]$, the probability $\mathrm{P}^{\pi}_{s_0}(\bar{A})$ in (9) can be expressed as

$$\mathrm{P}^{\pi}_{s_0}(\bar{A}) = E^{\pi}_{s_0}[\max_{k \in [0,N]} \mathbf{1}_{\bar{A}}(\mathbf{s}(k))]. \tag{13}$$

4.1. Probabilistic reachability computations

We next show how to compute $\mathrm{p}^{\pi}_{s_0}(A)$ and $\mathrm{P}^{\pi}_{s_0}(\bar{A})$ through a backward iterative procedure. To this purpose, we recall that a Markov policy $\pi \in \mathcal{M}_m$ is a sequence $\pi = (\mu_0, \mu_1, \mu_2, \ldots)$ of maps $\mu_l : \mathcal{S} \to \mathcal{U} \times \Sigma, l = 0, 1, 2, \ldots, N - 1$.

Multiplicative Cost. Define the set of functions $V^{\pi}_k : \mathcal{S} \to [0, 1]$, $k = 0, 1, \ldots, N$, as follows:

$$V^{\pi}_N(s) = \mathbf{1}_A(s)$$

$$V^{\pi}_k(s) = \mathbf{1}_A(s) \int_{\mathcal{S}^{N-k}} \prod_{l=k+1}^{N} \mathbf{1}_A(s_l) \prod_{l=k+1}^{N-1} T_s^{\mu_l}(\mathrm{d}s_{l+1}|s_l) \times T_s^{\mu_k}(\mathrm{d}s_{k+1}|s), \tag{14}$$

$k = 0, 1, \ldots, N - 1$, $s \in \mathcal{S}$, with $T_s^{\mu_h}(\cdot|s_h)$ standing for $T_s(\cdot|s_h, \mu_h(s_h))$. The maps $T_s^{\mu_h}(\cdot|s_h)$, $h = 0, 1, \ldots, N - 1$, are the one-step transition kernels of the embedded Markov process obtained by applying the Markov policy $\pi = (\mu_0, \mu_1, \ldots)$ to the DTSHS (see Section 2). Then, it is easily seen that, by (12), $V^{\pi}_0(s)$ evaluated at $s = s_0$ returns $\mathrm{p}^{\pi}_{s_0}(A)$ since $V^{\pi}_0(s) = E^{\pi}_s[\prod_{l=0}^{N} \mathbf{1}_A(\mathbf{s}(l))]$, $s \in \mathcal{S}$. Moreover, the probabilistic safe set with safety level $1 - \epsilon$, $\epsilon \in [0, 1]$, according to (6), can be expressed in terms of function V^{π}_0 as follows: $S^{\pi}(\epsilon) = \{s_0 \in \mathcal{S} : V^{\pi}_0(s_0) \geq 1 - \epsilon\}$.

Let $\mathcal{F}$ denote the set of functions from $\mathcal{S}$ to $\mathbb{R}$, and define the operator $H : \mathcal{S} \times \mathcal{U} \times \Sigma \times \mathcal{F} \to \mathbb{R}$ as follows:

$$H(s,(u,\sigma),Z) := T_q(q|s,u)\int_{\mathbb{R}^{n(q)}} Z((q,v))T_x(\mathrm{d}v|s,u) + \sum_{q'\neq q} T_q(q'|s,u)\int_{\mathbb{R}^{n(q')}} Z((q',v))R(\mathrm{d}v|s,\sigma,q'), \quad (15)$$

for any $s=(q,x)\in\mathcal{S}$, $(u,\sigma)\in\mathcal{U}\times\Sigma$, and $Z\in\mathcal{F}$. The operator H is easily seen to be a linear operator. Moreover, H applied to a constant function $\bar{Z}(s)=c$, $s\in\mathcal{S}$, returns the constant c for any value of the other arguments s and (u,σ): $H(s,(u,\sigma),\bar{Z})=c, \forall s\in\mathcal{S}$, $(u,\sigma)\in\mathcal{U}\times\Sigma$. This is because $H(s,(u,\sigma),Z)$ is the integral over $\mathcal{S}$ of function Z with respect to the (conditional) probability measure $T_s(\cdot|s,(u,\sigma))$ defined in (1).

With an argument inspired by a logic developed in Kumar and Varaiya (1986) for additive costs, we prove the following lemma.

Lemma 1. *Fix a Markov policy $\pi=(\mu_0,\mu_1,\ldots)\in\mathcal{M}_m$. The functions $V_k^\pi:\mathcal{S}\to[0,1]$, $k=0,1\ldots,N-1$, can be computed by the backward recursion:*

$$V_k^\pi(s)=\mathbf{1}_A(s)H(s,\mu_k(s),V_{k+1}^\pi),\quad s\in\mathcal{S}, \quad (16)$$

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

Proof. We start by observing that, given the definition of T_s in (1) in terms of its components, and that of H in (15), Eq. (16) can be rewritten as

$$V_k^\pi(s)=\mathbf{1}_A(s)\int_{\mathcal{S}} V_{k+1}^\pi(s_{k+1})T_s(\mathrm{d}s_{k+1}|s,\mu_k(s)).$$

From the expression in (14) of V_k^π, we have that

$$V_{N-1}^\pi(s) = \mathbf{1}_A(s)\int_{\mathcal{S}}\mathbf{1}_A(s_N)T_s^{\mu_{N-1}}(\mathrm{d}s_N|s) = \mathbf{1}_A(s)\int_{\mathcal{S}} V_N^\pi(s_N)T_s(\mathrm{d}s_N|s,\mu_{N-1}(s)),$$

so that Eq. (16) is proven for $k=N-1$. For $k<N-1$, V_k^π can be expanded as follows

$$V_k^\pi(s) = \mathbf{1}_A(s)\int_{\mathcal{S}}\mathbf{1}_A(s_{k+1})\left(\int_{\mathcal{S}^{N-k-1}}\prod_{l=k+2}^{N}\mathbf{1}_A(s_l)\times\prod_{l=k+2}^{N-1}T_s^{\mu_l}(\mathrm{d}s_{l+1}|s_l)T_s^{\mu_{k+1}}(\mathrm{d}s_{k+2}|s_{k+1})\right)T_s^{\mu_k}(\mathrm{d}s_{k+1}|s) = \mathbf{1}_A(s)\int_{\mathcal{S}}V_{k+1}^\pi(s_{k+1})T_s^{\mu_k}(\mathrm{d}s_{k+1}|s),$$

which concludes the proof. □

Max Cost. Define the set of functions $W_k^\pi:\mathcal{S}\to[0,1]$, $k=0,1,\ldots,N$, as follows:

$$W_N^\pi(s)=\mathbf{1}_{\bar{A}}(s)$$

$$W_k^\pi(s) = \mathbf{1}_{\bar{A}}(s)+\mathbf{1}_A(s)\int_{\mathcal{S}^{N-k}}\max_{l\in[k+1,N]}\mathbf{1}_{\bar{A}}(s_l)\times\prod_{l=k+1}^{N-1}T_s^{\mu_l}(\mathrm{d}s_{l+1}|s_l)T_s^{\mu_k}(\mathrm{d}s_{k+1}|s), \quad (17)$$

$k=0,1,\ldots,N-1$, $s\in\mathcal{S}$. Given the expression of $\mathrm{P}_{s_0}^\pi(\bar{A})$ as a max cost in (13), it is easy to show that $W_0^\pi(s)$ evaluated at $s=s_0$ returns $\mathrm{P}_{s_0}^\pi(\bar{A})$ since $W_0^\pi(s)=E_s^\pi[\max_{l\in[0,N]}\mathbf{1}_{\bar{A}}(\mathbf{s}(l))]$, $s\in\mathcal{S}$. Also, based on (10), the probabilistic safe set with safety level $1-\epsilon$, $\epsilon\in[0,1]$, can be expressed in terms of W_0^π as $S^\pi(\epsilon)=\{s_0\in\mathcal{S}: W_0^\pi(s_0)\leq\epsilon\}$. From the definition of W_k^π in (17), and that of H in (15), Lemma 2 follows.

Lemma 2. *Fix a Markov policy $\pi=(\mu_0,\mu_1,\ldots)\in\mathcal{M}_m$. The functions $W_k^\pi:\mathcal{S}\to[0,1]$, $k=0,1\ldots,N-1$, can be computed by the backward recursion:*

$$W_k^\pi(s)=\mathbf{1}_{\bar{A}}(s)+\mathbf{1}_A(s)H(s,\mu_k(s),W_{k+1}^\pi),\ s\in\mathcal{S}, \quad (18)$$

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

Proof. From the definition (17) of W_k^π, we have that

$$W_{N-1}^\pi(s) = \mathbf{1}_{\bar{A}}(s)+\mathbf{1}_A(s)\int_{\mathcal{S}}\mathbf{1}_{\bar{A}}(s_N)T_s^{\mu_{N-1}}(\mathrm{d}s_N|s) = \mathbf{1}_{\bar{A}}(s)+\mathbf{1}_A(s)\int_{\mathcal{S}}W_N^\pi(s_N)T_s^{\mu_{N-1}}(\mathrm{d}s_N|s)$$

so that Eq. (18) is proven for $k=N-1$. For $k<N-1$, W_k^π can be expanded as follows

$$W_k^\pi(s) = \mathbf{1}_{\bar{A}}(s)+\mathbf{1}_A(s)\int_{\mathcal{S}}\Big(\mathbf{1}_{\bar{A}}(s_{k+1})+\mathbf{1}_A(s_{k+1})\times\int_{\mathcal{S}^{N-k-1}}\max_{l\in[k+2,N]}\mathbf{1}_{\bar{A}}(s_l)\prod_{l=k+2}^{N-1}T_s^{\mu_l}(\mathrm{d}s_{l+1}|s_l)\times T_s^{\mu_{k+1}}(\mathrm{d}s_{k+2}|s_{k+1})\Big)T_s^{\mu_k}(\mathrm{d}s_{k+1}|s) = \mathbf{1}_{\bar{A}}(s)+\mathbf{1}_A(s)\int_{\mathcal{S}}W_{k+1}^\pi(s_{k+1})T_s^{\mu_k}(\mathrm{d}s_{k+1}|s)$$

which concludes the proof. □

It is worth noting that the iterative backward recursion derived in Lemma 2 is similar to that in Prandini and Hu (2006a,b) for reachability computations on the Markov chain approximation of certain classes of uncontrolled continuous time stochastic hybrid systems.

Equivalence of the two representations. Since for any sequence of state values $s_l\in\mathcal{S}, l=0,1,\ldots,N$, $\prod_{l=k}^N\mathbf{1}_A(s_l)=1-\max_{l\in[k,N]}\mathbf{1}_{\bar{A}}(s_l)$, $k=0,1,\ldots,N$, not surprisingly the following equivalence result holds (see Fig. 1).

Lemma 3. *Fix a Markov policy $\pi=(\mu_0,\mu_1,\ldots)\in\mathcal{M}_m$. Then, for any $k\in[0,N]$, $W_k^\pi(s)=1-V_k^\pi(s)$, $s\in\mathcal{S}$.* □

Proof. The statement trivially holds for $k=N$. Assume now that it holds at time $k+1$. Then,

$$\begin{aligned} W_k^\pi(s) &= [\text{by Lemma 2}]\\ &= \mathbf{1}_{\bar{A}}(s)+\mathbf{1}_A(s)H(s,\mu_k(s),W_{k+1}^\pi)\\ &= [\text{by induction assumption}]\\ &= 1-\mathbf{1}_A(s)+\mathbf{1}_A(s)H(s,\mu_k(s),1-V_{k+1}^\pi)\\ &= [\text{by the properties of the operator } H]\\ &= 1-\mathbf{1}_A(s)+\mathbf{1}_A(s)(1-H(s,\mu_k(s),V_{k+1}^\pi))\\ &= 1-\mathbf{1}_A(s)H(s,\mu_k(s),V_{k+1}^\pi)\\ &= [\text{by Lemma 1}]\\ &= 1-V_k^\pi(s),\quad s\in\mathcal{S}, \end{aligned}$$

so that the statement holds for any $k=0,1,\ldots,N$. □

4.2. Maximal probabilistic safe sets computation

The calculation of the maximal probabilistic safe set $S^\star(\epsilon)$ defined in (7) amounts to finding the supremum over the Markov policies of the probability $\mathrm{p}_{s_0}^\pi(A)$ of remaining within the safe set A starting from s_0, for all s_0 inside A (the probability of remaining within A starting from $s_0\notin A$ is 0 for any policy). A policy that achieves this supremum is said to be *maximally safe*. More precisely,

Definition 4 (*Maximally Safe Policy*). Let $\mathcal{H}$ be a DTSHS, and $A\in\mathcal{B}(\mathcal{S})$ a safe set. A Markov policy $\pi^\star$ is maximally safe if $\mathrm{p}_{s_0}^{\pi^\star}(A)=\sup_{\pi\in\mathcal{M}_m}\mathrm{p}_{s_0}^\pi(A), \forall s_0\in A$. □

Note that, in view of Lemma 3, a maximally safe policy can be equivalently characterized as that policy $\pi^\star \in \mathcal{M}_m$ that achieves the minimum over A of $\mathrm{P}^\pi_{s_0}(\bar{A})$: $\mathrm{P}^{\pi^\star}_{s_0}(\bar{A}) = \inf_{\pi\in\mathcal{M}_m} \mathrm{P}^\pi_{s_0}(\bar{A}), \forall s_0 \in A$.

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, and describe an algorithm to compute $\sup_{\pi\in\mathcal{M}_m} \mathrm{p}^\pi_{s_0}(A)$ in terms of the multiplicative cost, and $\inf_{\pi\in\mathcal{M}_m} \mathrm{P}^\pi_{s_0}(\bar{A})$ in terms of the max cost.

Multiplicative cost. We now show how to compute a maximally safe Markov policy $\pi^\star \in \mathcal{M}_m$ through a recursion similar to that in Lemma 1, based on the representation (12) of $\mathrm{p}^\pi_{s_0}(A)$ as a multiplicative cost. The proof is inspired by Bertsekas and Shreve (1996, Section 11.3) addressing a finite horizon stochastic optimal control problem with a multiplicative cost to be minimized.

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

$$V^\star_k(s) = \sup_{(u,\sigma)\in\mathcal{U}\times\Sigma} \mathbf{1}_A(s) H(s,(u,\sigma),V^\star_{k+1}), \tag{19}$$

$s \in \mathcal{S}$, initialized with $V^\star_N(s) = \mathbf{1}_A(s)$, $s \in \mathcal{S}$.

Then, $V^\star_0(s_0) = \sup_{\pi\in\mathcal{M}_m} \mathrm{p}^\pi_{s_0}(A)$, $s_0 \in \mathcal{S}$.

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

$$\mu^\star_k(s) \in \arg \sup_{(u,\sigma)\in\mathcal{U}\times\Sigma} H(s,(u,\sigma),V^\star_{k+1}), \quad s \in A, \tag{20}$$

then, $\pi^\star = (\mu^\star_0, \mu^\star_1, \ldots, \mu^\star_{N-1})$ is a maximally safe Markov policy. A sufficient condition for the existence of such a $\pi^\star$ is that $U_k(s,\lambda) = \{(u,\sigma) \in \mathcal{U}\times\Sigma : H(s,(u,\sigma),V^\star_{k+1}) \geq \lambda\}$ is compact for all $s \in A$, $\lambda \in \mathbb{R}$, $k \in [0, N-1]$. □

Proof. For ease of reference to Bertsekas and Shreve (1996, Section 11.3), we set $J^\pi_k := -V^\pi_{N-k}$, $\pi \in \mathcal{M}_m$, and $J^\star_k := -V^\star_{N-k}$, $k = 0,1,\ldots,N$, and rewrite Eq. (20) and the recursions (16) and (19) in terms of these functions as:

$$\mu^\star_k(s) \in \arg \inf_{(u,\sigma)\in\mathcal{U}\times\Sigma} H(s,(u,\sigma),J^\star_{N-k-1}), \quad s \in A, \tag{21}$$

$$J^\pi_k(s) = \mathbf{1}_A(s) H(s,\mu_{N-k}(s),J^\pi_{k-1}) \tag{22}$$

$$J^\star_k(s) = \inf_{(u,\sigma)\in\mathcal{U}\times\Sigma} \mathbf{1}_A(s) H(s,(u,\sigma),J^\star_{k-1}), \tag{23}$$

initialized with $J^\pi_0(s) = J^\star_0(s) = -\mathbf{1}_A(s)$, $s \in \mathcal{S}$. Notice that the inclusion signs in (20) and (21) indicate the possible non-uniqueness of the optimal policy.

Consider a (universally measurable) function $\mu : \mathcal{S} \to \mathcal{U}\times\Sigma$ and define the map $T_\mu : \mathcal{F} \to \mathcal{F}$ as $T_\mu[J](s) = K(s,\mu(s),J)$, $s \in \mathcal{S}$, where $K(s,(u,\sigma),J) = \mathbf{1}_A(s) H(s,(u,\sigma),J)$, $s \in \mathcal{S}$, $(u,\sigma) \in \mathcal{U}\times\Sigma$, $J \in \mathcal{F}$.

Let $\tilde{\mathcal{F}} \subset \mathcal{F}$ denote the set of universally measurable real functions $J : \mathcal{S} \to \mathbb{R}$. The map T_μ preserves the universal measurability property: if $J \in \tilde{\mathcal{F}}$, then, $T_\mu[J] \in \tilde{\mathcal{F}}$. This is because the integration of a universally measurable function with respect to the stochastic kernel involved in the computation of $H(s,\mu(s),J)$ (see (15)) is a universally measurable function, and its product with the Borel measurable function $\mathbf{1}_A(s)$ remains universally measurable. Observe that, since the recursion (22) can be rewritten as $J^\pi_k = T_{\mu_{N-k}}[J^\pi_{k-1}]$ and by definition $J^\pi_0 \in \tilde{\mathcal{F}}$, we then have that $J^\pi_k \in \tilde{\mathcal{F}}$, $k = 1,2,\ldots,N$.

The map T_μ also satisfies the following properties: for all $J, J' \in \tilde{\mathcal{F}}$ such that $J(s) \leq J'(s)$, $\forall s \in \mathcal{S}$, then $T_\mu[J](s) \leq T_\mu[J'](s)$, $\forall s \in \mathcal{S}$ (monotonicity, (Bertsekas & Shreve, 1996, Section 6.2)), and for any $J \in \tilde{\mathcal{F}}$ and any real number $r > 0$,

$$T_\mu[J](s) \leq T_\mu[J+r](s) \leq T_\mu[J](s) + r, \quad s \in \mathcal{S}. \tag{24}$$

The monotonicity property immediately follows from the definition of T_μ. As for property (24), it is easily shown observing that, by the definition of K and the properties of the following chain of equalities holds: $K(s,(u,\sigma),J+r) = \mathbf{1}_A(s)H(s,(u,\sigma),J+r) = \mathbf{1}_A(s)H(s,(u,\sigma),J) + \mathbf{1}_A(s)r$, $s \in \mathcal{S}$, $(u,\sigma) \in \mathcal{U}\times\Sigma$, and, hence, given that $\mathbf{1}_A(s)$ is either equal to 0 or to 1, $K(s,(u,\sigma),J) \leq K(s,(u,\sigma),J+r) \leq K(s,(u,\sigma),J) + r$, $s \in \mathcal{S}$, $(u,\sigma) \in \mathcal{U}\times\Sigma$.

Now, define the map $T : \mathcal{F} \to \mathcal{F}$ as $T[J](s) = \inf_{(u,\sigma)\in\mathcal{U}\times\Sigma} K(s,(u,\sigma),J)$, $s \in \mathcal{S}$. Then, the recursion (23) can be rewritten as $J^\star_k = T[J^\star_{k-1}]$, and, from this latter expression, it follows that $J^\star_k = T^k[J^\star_0]$, $k = 0,1,\ldots,N$, where $T^0[J] = J$ and $T^k[J] = T[T^{k-1}[J]]$. Let $\mathcal{F}^* \subset \tilde{\mathcal{F}}$ denote the set of lower-semianalytic functions. The map T preserves the lower-semianalytic property: if $J \in \mathcal{F}^*$, then, $T[J] \in \mathcal{F}^*$. This follows from Bertsekas and Shreve (1996, Proposition 7.47), given that $H(s,(u,\sigma),J)$ is lower-semianalytic as a function of its arguments $s \in \mathcal{S}$ and $(u,\sigma) \in \mathcal{U}\times\Sigma$ (Bertsekas & Shreve, 1996, Proposition 7.48), and, hence, $K(s,(u,\sigma),J) = \mathbf{1}_A(s)H(s,(u,\sigma),J)$ is lower-semianalytic as well (Bertsekas & Shreve, 1996, Lemma 7.30(4)). Since $J^\star_k = T[J^\star_{k-1}]$ and $J^\star_0 \in \mathcal{F}^*$, we then have that $J^\star_k \in \mathcal{F}^*$, $k = 1,2,\ldots,N$.

After these preliminary considerations, we prove by induction on the horizon N the following two statements:

1. $$\inf_\pi E^\pi_s\left[-\prod_{k=0}^N \mathbf{1}_A(s_k)\right] = T^N[J^\star_0](s), \quad s \in \mathcal{S}$$
2. $\forall \epsilon > 0$, $\exists \pi_\epsilon = (\mu_{\epsilon,0}, \mu_{\epsilon,1}, \ldots) \in \mathcal{M}_m : \forall s \in \mathcal{S}$,
$$\inf_\pi E^\pi_s\left[-\prod_{k=0}^N \mathbf{1}_A(s_k)\right] \leq J^{\pi_\epsilon}_N(s) \leq \inf_\pi E^\pi_s\left[-\prod_{k=0}^N \mathbf{1}_A(s_k)\right] + \epsilon.$$

Note that by the first statement, it follows that $V^\star_0(s_0) = -J^\star_N(s_0) = -T^N[J^\star_0](s_0) = \sup_{\pi\in\mathcal{M}_m} \mathrm{p}^\pi_{s_0}(A)$, for any $s_0 \in \mathcal{S}$, so that the first part of the theorem is proven. As for the second statement, observe that, for any $\pi_\epsilon \in \mathcal{M}_m$

$$\begin{aligned} J^{\pi_\epsilon}_N(s) &= -V^{\pi_\epsilon}_0(s) = E^{\pi_\epsilon}_s\left[-\prod_{k=0}^N \mathbf{1}_A(s_k)\right] \\ &\geq \inf_\pi E^\pi_s\left[-\prod_{k=0}^N \mathbf{1}_A(s_k)\right], \end{aligned}$$

so that the part of the second statement that needs to be actually proven is the right-hand-side.

The statements clearly hold for $N = 0$. Suppose that they hold for $N = h$. This implies that $\forall \epsilon > 0, \exists \pi_\epsilon = (\mu_{\epsilon,0}, \mu_{\epsilon,1}, \ldots) \in \mathcal{M}_m$ such that

$$J^{\pi_\epsilon}_h(s) \leq \inf_\pi E^\pi_s\left[-\prod_{l=0}^h \mathbf{1}_A(s_l)\right] + \epsilon, \quad s \in \mathcal{S}.$$

For any universally measurable function $\mu : \mathcal{S} \to \mathcal{U}\times\Sigma$, we then have that, by the monotonicity of T_μ and by (24),

$$\begin{aligned} T_\mu[J^{\pi_\epsilon}_h](s) &\leq T_\mu\left[\inf_\pi E^\pi_s\left[-\prod_{l=0}^h \mathbf{1}_A(s_l)\right] + \epsilon\right] \\ &\leq T_\mu\left[\inf_\pi E^\pi_s\left[-\prod_{l=0}^h \mathbf{1}_A(s_l)\right]\right] + \epsilon, \quad s \in \mathcal{S}. \end{aligned}$$

Now, if we consider policy $\bar{\pi}_\epsilon = (\mu, \mu_{\epsilon,0}, \mu_{\epsilon,1}, \ldots)$, it is easily seen that

$$\inf_\pi E^\pi_s\left[-\prod_{l=0}^{h+1} \mathbf{1}_A(s_l)\right] \leq J^{\bar{\pi}_\epsilon}_{h+1}(s) = T_\mu[J^{\pi_\epsilon}_h](s), \quad s \in \mathcal{S},$$

which, combined with the inequality above, leads to:

$$\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h+1}\mathbf{1}_A(s_l)\right]\le T_\mu\left[\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h}\mathbf{1}_A(s_l)\right]\right]+\epsilon,\quad s\in\mathcal{S},$$

for any universally measurable function $\mu:\mathcal{S}\to\mathcal{U}\times\Sigma$. From this, it follows that

$$\begin{aligned}\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h+1}\mathbf{1}_A(s_l)\right]&\le T\left[\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h}\mathbf{1}_A(s_l)\right]\right]\\&=T^{h+1}[J_0^\star](s),\quad s\in\mathcal{S},\end{aligned}$$

where the last equality is due to the induction hypothesis. On the other hand, we clearly have that

$$T^{h+1}[J_0^\star](s)\le\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h+1}\mathbf{1}_A(s_l)\right],\quad s\in\mathcal{S},$$

which allows to conclude that

$$\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h+1}\mathbf{1}_A(s_l)\right]=T^{h+1}[J_0^\star](s),\quad s\in\mathcal{S}.\tag{25}$$

Let us now proceed with the second statement.

By the induction hypothesis, for any $\bar\epsilon>0$ there exists a $\bar\pi=(\bar\mu_0,\bar\mu_1,\ldots)\in\mathcal{M}_m$ such that

$$J_h^{\bar\pi}(s)\le\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h}\mathbf{1}_A(s_l)\right]+\frac{\bar\epsilon}{2},\quad s\in\mathcal{S}.$$

Also, by Bertsekas and Shreve (1996, Proposition 7.50) there exists a universally measurable function $\bar\mu:\mathcal{S}\to\mathcal{U}\times\Sigma$ such that

$$T_{\bar\mu}\left[E_s^\pi\left[-\prod_{l=0}^{h}\mathbf{1}_A(s_l)\right]\right](s)\le T\left[E_s^\pi\left[-\prod_{l=0}^{h}\mathbf{1}_A(s_l)\right]\right](s)+\frac{\bar\epsilon}{2},$$

$s\in\mathcal{S}$. Then, if we consider $\pi_{\bar\epsilon}=(\bar\mu,\bar\mu_0,\bar\mu_1,\ldots)$, by the monotonicity of $T_{\bar\mu}$ and (24), we obtain

$$\begin{aligned}J_{h+1}^{\pi_{\bar\epsilon}}(s)=T_{\bar\mu}[J_h^{\bar\pi}](s)&\le T_{\bar\mu}\left[\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h}\mathbf{1}_A(s_l)\right]\right]+\frac{\bar\epsilon}{2}\\&\le T\left[\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h}\mathbf{1}_A(s_l)\right]\right]+\bar\epsilon,\quad s\in\mathcal{S}.\end{aligned}$$

By the induction hypothesis and (25), we finally get

$$J_{h+1}^{\pi_{\bar\epsilon}}(s)\le T^{h+1}[J_0^\star]+\bar\epsilon=\inf_\pi E_s^\pi\left[-\prod_{l=0}^{h+1}\mathbf{1}_A(s_l)\right]+\bar\epsilon,$$

$s\in\mathcal{S}$, which concludes the proof of the two statements.

Next, we show that $\pi^\star=(\mu_0^\star,\mu_1^\star,\ldots)$ satisfying (21) is a Markov policy and that it is maximally safe. To this purpose, note first that a function $\mu_k^\star$ satisfying (21) can be characterized through the equation

$$\begin{aligned}T_{\mu_k^\star}[J_{N-k-1}^\star](s)&=\inf_{(u,\sigma)\in\mathcal{U}\times\Sigma}K(s,(u,\sigma),J_{N-k-1}^\star)\\&=J_{N-k}^\star(s),\quad s\in\mathcal{S}.\end{aligned}$$

As discussed at the beginning of this proof, $J_{N-k}^\star\in\mathcal{F}^*$ and, hence, $K(s,(u,\sigma),J_{N-k}^\star)$ is lower-semianalytic as a function of $s\in\mathcal{S}$ and $(u,\sigma)\in\mathcal{U}\times\Sigma$. Then, if its infimum with respect to $(u,\sigma)\in\mathcal{U}\times\Sigma$ is attained for any $s\in\mathcal{S}$ (for $s\in\bar A$ this is always the case, given the sufficient condition), the resulting function $\mu_k^\star:\mathcal{S}\to\mathcal{U}\times\Sigma$ is universally measurable, (Bertsekas & Shreve, 1996, Proposition 7.50). Now observe that

$$\begin{aligned}&\inf_\pi E_s^\pi\left[-\prod_{l=0}^{N}\mathbf{1}_A(s_l)\right]=J_N^\star(s)\\&=T_{\mu_0^\star}[J_{N-1}^\star](s)=T_{\mu_0^\star}[T_{\mu_1^\star}[J_{N-2}^\star]](s)\\&=\cdots=T_{\mu_0^\star}[T_{\mu_1^\star}[\ldots T_{\mu_{N-1}^\star}[J_0^\star]]](s)\\&=J_N^{\pi^\star}(s)=E_s^{\pi^\star}\left[-\prod_{l=0}^{N}\mathbf{1}_A(s_l)\right],\end{aligned}$$

$s\in\mathcal{S}$, which shows that $\pi^\star$ is maximally safe.

For any $s\in\mathcal{S}$ and $k\in[0,N-1]$, a sufficient condition for the existence of a minimum over $\mathcal{U}\times\Sigma$ of function $K(s,(u,\sigma),J_{N-k-1}^\star)$ is that $Z_k(s,\alpha)=\{(u,\sigma)\in\mathcal{U}\times\Sigma:K(s,(u,\sigma),J_{N-k-1}^\star)\le\alpha\}$ is compact, (Bertsekas & Shreve, 1996, Lemma 3.1). Since $J_{N-k-1}^\star=-V_{k+1}^\star$, then $K(s,(u,\sigma),J_{N-k-1}^\star)=\mathbf{1}_A(s)H(s,(u,\sigma),J_{N-k-1}^\star)=-\mathbf{1}_A(s)H(s,(u,\sigma),V_{k+1}^\star)$, from which the condition on $U_k(s,\lambda)$ easily follows. □

Max cost. In the following theorem, we describe an algorithm to compute a maximally safe Markov policy $\pi^\star\in\mathcal{M}_m$ based on the representation (13) of $\mathrm{P}_{s_0}^\pi(\bar A)$ as a max cost, by a recursion similar to that in Lemma 2.

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

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

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

Then, $W_0^\star(s_0)=\inf_{\pi\in\mathcal{M}_m}\mathrm{P}_{s_0}^\pi(\bar A)$, $s_0\in\mathcal{S}$.

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

$$\mu_k^\star(s)\in\arg\inf_{(u,\sigma)\in\mathcal{U}\times\Sigma}H(s,(u,\sigma),W_{k+1}^\star),\quad\forall s\in A,$$

then, $\pi^\star=(\mu_0^\star,\mu_1^\star,\ldots,\mu_{N-1}^\star)$ is a maximally safe Markov policy. A sufficient condition for the existence of such a $\pi^\star$ is that $U_k(s,\lambda)=\{(u,\sigma)\in\mathcal{U}\times\Sigma:H(s,(u,\sigma),W_{k+1}^\star)\le\lambda\}$ is compact for all $s\in A$, $\lambda\in\mathbb{R}$, $k\in[0,N-1]$. □

Proof. We start proving that for any $k\in[0,N]$,

$$W_k^\star(s)=1-V_k^\star(s),\quad s\in\mathcal{S}.\tag{26}$$

The statement is trivially satisfied for $k=N$, since $W_N^\star(s)=\mathbf{1}_{\bar A}(s)=1-\mathbf{1}_A(s)=1-V_N^\star(s)$, $s\in\mathcal{S}$. Assume that it is valid for $k+1$. Then,

$$\begin{aligned}W_k^\star(s)&=\inf_{(u,\sigma)\in\mathcal{U}\times\Sigma}(\mathbf{1}_{\bar A}(s)+\mathbf{1}_A(s)H(s,(u,\sigma),W_{k+1}^\star))\\&=\text{[by induction assumption]}\\&=\inf_{(u,\sigma)\in\mathcal{U}\times\Sigma}(\mathbf{1}_{\bar A}(s)+\mathbf{1}_A(s)H(s,(u,\sigma),1-V_{k+1}^\star))\\&=\text{[by the properties of the operator }H\text{]}\\&=\inf_{(u,\sigma)\in\mathcal{U}\times\Sigma}(\mathbf{1}_{\bar A}(s)+\mathbf{1}_A(s)(1-H(s,(u,\sigma),V_{k+1}^\star)))\\&=1-\sup_{(u,\sigma)\in\mathcal{U}\times\Sigma}\mathbf{1}_A(s)H(s,(u,\sigma),V_{k+1}^\star)\\&=\text{[by Theorem 1]}\\&=1-V_k^\star(s),\quad s\in\mathcal{S}.\end{aligned}$$

It then easily follows from Theorem 1 and the definitions of $\mathrm{p}_{s_0}^\pi(A)$ and $\mathrm{P}_{s_0}^\pi(\bar A)$ that

$$\begin{aligned}W_0^\star(s_0)&=1-V_0^\star(s_0)=1-\sup_{\pi\in\mathcal{M}_m}\mathrm{p}_{s_0}^\pi(A)\\&=\inf_{\pi\in\mathcal{M}_m}(1-\mathrm{p}_{s_0}^\pi(A))=\inf_{\pi\in\mathcal{M}_m}\mathrm{P}_{s_0}^\pi(\bar A).\end{aligned}$$

Furthermore, in view of the duality equation (26) the characterization through the $V_k^\star$ functions of a maximally safe policy $\mu_k^\star : \mathcal{S} \to \mathcal{U} \times \Sigma, k \in [0, N-1]$, in Theorem 1 can be equivalently expressed in terms of the $W_k^\star$ functions as follows:

$$\mu_k^\star(s) \in \arg \sup_{(u,\sigma)\in\mathcal{U}\times\Sigma} H(s,(u,\sigma),V_{k+1}^\star);$$

$$\mu_k^\star(s) \in \arg \sup_{(u,\sigma)\in\mathcal{U}\times\Sigma} H(s,(u,\sigma),1-W_{k+1}^\star);$$

$$\mu_k^\star(s) \in \arg \inf_{(u,\sigma)\in\mathcal{U}\times\Sigma} H(s,(u,\sigma),W_{k+1}^\star), \quad s \in A.$$

A sufficient condition on the control space to ensure the existence of this optimal argument at each time step is again easily derived from the corresponding one in Theorem 1. □

Remark 1. If the control input spaces $\mathcal{U}$ and Σ are both finite sets, then a maximally safe policy is guaranteed to exist. □

5. Case study: Temperature regulation — control

We address the temperature regulation problem described in Section 3.

The objective is to maintain the temperature of r rooms within a certain range over some finite time horizon by heating one room at a time. To this purpose we devise a Markov policy that decides at each time instant which room should be heated based on the current value of the temperature in the r rooms. This control design problem can be reformulated as a safety problem. The 'safe' set is represented here by the desired temperature range for any discrete state.

We present the results obtained in the $r = 2$ rooms case. The temperature is measured in degrees Celsius and one discrete time unit corresponds to $\Delta = 10$ min. The discrete time horizon is $[0, N]$ with $N = 60$, which corresponds to an effective length of $N\Delta = 600$ min.

The discrete state space is $\mathcal{Q} = \{\mathtt{ON}_1, \mathtt{ON}_2, \mathtt{OFF}\}$ and the continuous state space is $\mathbb{R}^2$ in each mode $q \in \mathcal{Q}$. The desired temperature interval is $[17.5, 22]$ in both rooms. Thus, the safe set A is given by $A = \mathcal{Q} \times A_x$ with $A_x := [17.5, 22] \times [17.5, 22]$. The parameters values in Eq. (2) are set equal to: $x_a = 6$, $b_1 = b_2 = 0.0167$, $a_{12} = a_{21} = 0.022$, $c_1 = 0.8$, $c_2 = 0.9333$, and $\nu^2 = 0.06$.

The transition control input takes on values in $\mathcal{U} = \{\mathtt{SW}_1, \mathtt{SW}_2, \mathtt{SW}_{\mathtt{OFF}}\}$ and affects the evolution of the controlled Markov chain governing the discrete transitions of the DTSHS model. We suppose that when a command aimed at commuting from a mode to a different one is issued, then the prescribed transition actually occurs with a probability 0.8, whereas with probability 0.1 the situation remains unchanged (which models a delay) and with probability 0.1 a transition to the third, non-recommended node, occurs (which models a faulty behavior). Instead, when a command of remaining in the current node is issued, this happens with probability 1. These specifications can be precisely formalized by appropriately defining the controlled Markov chain transition probabilities $\{\alpha_{qq'}(u), q, q' \in \mathcal{Q}\}$ introduced in (4), for any $u \in \mathcal{U}$. For instance, for $u = \mathtt{SW}_1$, $\alpha_{\mathtt{ON}_1\mathtt{ON}_1}(\mathtt{SW}_1) = 1$, $\alpha_{\mathtt{ON}_2\mathtt{ON}_1}(\mathtt{SW}_1) = 0.8$, $\alpha_{\mathtt{ON}_2\mathtt{ON}_2}(\mathtt{SW}_1) = 0.1$, $\alpha_{\mathtt{OFF}\mathtt{ON}_1}(\mathtt{SW}_1) = 0.8$, and $\alpha_{\mathtt{OFF}\mathtt{OFF}}(\mathtt{SW}_1) = 0.1$, the other probabilities $\alpha_{qq'}(\mathtt{SW}_1)$ being determined by the normalization condition $\sum_{q'\in\mathcal{Q}} \alpha_{qq'}(\mathtt{SW}_1) = 1, q \in \mathcal{Q}$.

The dynamic programming recursion described in Section 4.2 can be used to compute a maximally safe policy $\pi^\star = (\mu_0^\star, \mu_1^\star, \ldots, \mu_{N-1}^\star)$, $\mu_k^\star : \mathcal{S} \to \mathcal{U}$, $k = 0, 1, \ldots, N-1$, and the maximal probabilistic safe sets $S^\star(\epsilon)$, $\epsilon \in [0, 1]$. In the implementation, the multiplicative setup was chosen, and computations were performed in MATLAB.

Fig. 2 shows some 'optimal' sample paths of the continuous state component of the DTSHS executions associated with the maximally safe policy $\pi^\star$ and different initial conditions. The initial conditions were chosen at random, according to the uniform distribution, over the safe set A. Note that, as expected, most of the temperature sample paths tend toward the central area of A_x, and those sample paths that exit A_x correspond to initial conditions close to its boundary. This is due partly to the delay in the commutations, and partly to the noise affecting the temperature dynamics.

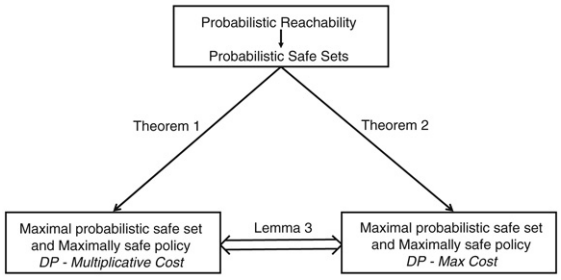

Fig. 1. Dual interpretation of the probabilistic reachability and safety problem.

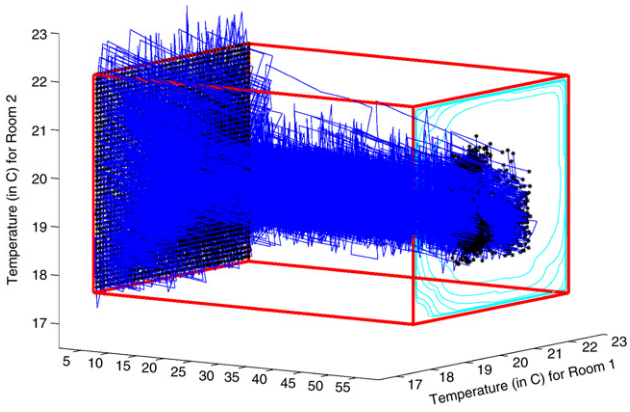

Fig. 2. Sample paths of the two rooms temperatures for executions corresponding to different initial conditions, under the same maximally safe policy.

In Fig. 3, we represent the component of the maximal probabilistic safe set $S^\star(\epsilon)$ associated with the discrete state OFF, that is $\{x \in \mathbb{R}^2 : (\mathtt{OFF}, x) \in S^\star(\epsilon)\} \subseteq A_x$, for different safety levels $1-\epsilon$. The plots corresponding to the discrete modes $\mathtt{ON}_1$ and $\mathtt{ON}_2$ are similar. As expected, the maximal probabilistic safe sets get smaller as the required safety level $1-\epsilon$ grows. Also, their shape reveals some asymmetry due to the structure of the temperature dynamics. Because of the low value of the ambient temperature ($x_a = 6$), the temperature tends naturally to decrease (see Eq. (2)).

The values taken by function $\mu_0^\star : \mathcal{S} \to \mathcal{U}$ over the set A_x when $q = \mathtt{OFF}$ are plotted in Fig. 4. $\mu_0^\star(\mathtt{OFF}, x)$ is the maximally safe transition control input issued at time $k = 0$ when $\mathbf{s}(0) = (\mathtt{OFF}, x)$. The maximally safe controls for the other time steps k within the horizon $[0, N]$ are indeed really similar to the one in Fig. 4, except for the very final time steps in the interval $[0, N]$. This means that, in practice, the *stationary Markov policy* with $\mu_k = \mu_0^\star$, $k = 0, 1, \ldots, N-1$, is nearly maximally safe. The interested reader is referred to a related, lower dimensional, case study reported in Abate, Amin, Prandini, Lygeros, and Sastry (2006), where this issue of nearly-stationarity is addressed in details.

6. Extensions

The described approach to stochastic reachability can be extended in several directions to address different interesting analysis and control design problems.

Other approaches in the literature, which focus on the infinite time horizon case and resort to the solution of a continuous-time

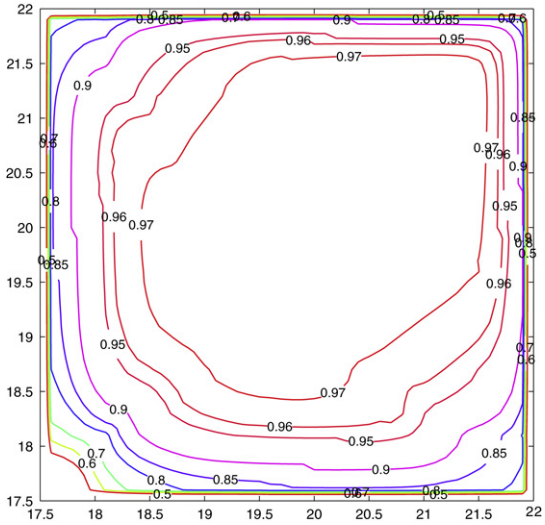

Fig. 3. Maximal probabilistic safe sets corresponding to different safety levels (0.5, 0.6, 0.7, 0.8, 0.85, 0.9, 0.95,0.96, and 0.97) within the discrete state OFF. The temperature of room 1 is reported on the horizontal axis, and that of room 2 on the vertical axis.

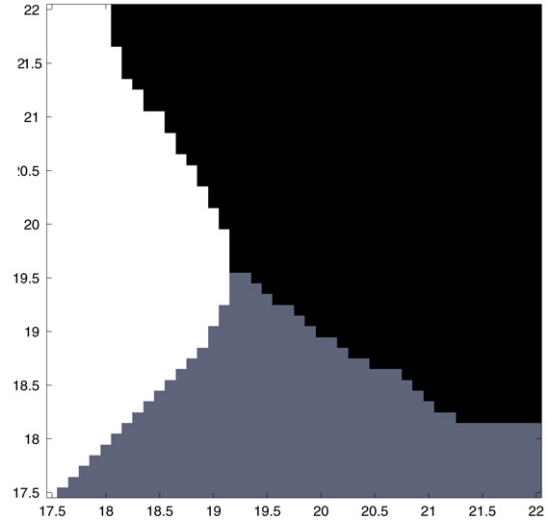

Fig. 4. Value taken by the maximally safe transition input at time $k = 0$ over the set of desired temperatures, when the heating system is in the OFF mode. The temperature of room 1 is reported on the horizontal axis, and that of room 2 on the vertical axis. The colors *black*, *white*, and *grey* respectively stand for the transition input command $\mathtt{SW}_{\mathtt{OFF}}$, $\mathtt{SW}_1$, and $\mathtt{SW}_2$.

dynamic programming equation, (Mitchell & Templeton, 2005), are structurally constrained to consider a fixed set. In contrast, in the present work the problem of evaluating the probability that the state of the stochastic hybrid system will remain within a *time-varying set* during some finite time horizon can be easily addressed. The dynamic programming equations in Section 4 should be simply adapted by considering at each iteration the corresponding set, and initializing the recursion with the appropriate indicator function. More specifically, given a safe set $A(k) \in \mathcal{B}(\mathcal{S})$, $k = 0, 1, \ldots, N$, the backwards recursions appearing in Lemma 1 and Theorem 1 should be changed by considering at each iteration k the corresponding set $A(k)$. Similarly for Lemma 2 and Theorem 2, where sets $\bar{A}(k)$ should be used.

Some problems arising in *regulation theory* can be reformulated within a reachability framework (Bertsekas, 1972). Suppose that the aim is to steer the state of the system close to some operating condition and that the requirements are such that one can introduce a small neighborhood around such a condition and a time-varying region shrinking to that neighborhood, reflecting admissible deviations from the desired system behavior in the transient. Then, the regulation problem can be rephrased as a safety problem where the 'safe' set is represented by the introduced time-varying region. The interested reader is referred to Abate et al. (2006) for an example of application of the proposed reachability methodology to one such a regulation problem.

By extending the approach to stochastic reachability to the *infinite time horizon* case, (Abate et al., 2006), *practical stabilization* problems, (Picasso & Bicchi, 2005), could be addressed. If the state of the system has to be driven to a certain neighborhood of an operating point and maintained there indefinitely, then one can think of structuring the problem in two stages: a first stage where a finite horizon time-varying reachability problem is solved (regulation part) and a second stage where an infinite horizon time-invariant problem is solved (invariance part).

More challenging extensions of the approach include the treatment of *optimal control problems with safety constraints*, as in Lygeros et al. (1999). It is in fact quite common when the controlled system is required to behave optimally according to some performance criterion, while evolving within a safe/desired set. In some cases, for example when the system performance is evaluated in terms of an additive cost, a relaxed version of the optimal constrained control problem can be formulated where a new state component representing at each time instant the cumulative cost is introduced and the objective is to maintain the system within an extended safe region where the cost is sufficiently low. As an example, suppose that in the considered temperature regulation case study we want to limit the number of transitions. We could then assign a unitary cost to each commutation, add a state variable counting the number of commutations, and keep this new variable within a bounded region around the origin, with an upper limit corresponding to the total number of allowed commutations. Clearly, the caveat to this approach is that an increase in dimensionality has to be taken in consideration as a tradeoff.

7. Conclusions

In this paper, we introduced a model for controlled discrete time stochastic hybrid systems. With reference to such a model, we formalized the finite-horizon stochastic reachability problem, which consists in maintaining the state of the system within a given safe set over some finite time horizon. We showed how the problem can be solved by designing an appropriate Markov policy through two equivalent dynamic programming schemes. For illustrative purposes, the problem of maintaining the temperature of two rooms within a desired range was considered as a case study.

Promising future directions for this research, from both a theoretical and an application-oriented viewpoint, were outlined in Section 6.

Acknowledgement

Thanks to Saurabh Amin for his contributions to the project.

References

Abate, A., Amin, S., Prandini, M., Lygeros, J., & Sastry, S. (2006). Probabilistic reachability and safe sets computation for discrete time stochastic hybrid systems. In *Proceedings of the 45th IEEE conference on decision and control* (pp. 258–263).

Alur, R., Henzinger, T., Lafferriere, G., & Pappas, G. J. (2000). Discrete abstractions of hybrid systems. *Proceedings of the IEEE*, *88*(2), 971–984.

Asarin, E., Bournez, O., Dang, T., & Maler, O. (2000). Approximate reachability analysis of piecewise linear dynamical systems. In N. Lynch, & B. Krogh (Eds.), *Lecture notes in computer science*: *Vol. 1790. Hybrid systems: Computation and control* (pp. 21–31). Springer Verlag.

Balluchi, A., Benvenuti, L., Di Benedetto, M. D., Miconi, G. M., Pozzi, U., Villa, T., et al. (2000). Maximal safe set computation for idle speed control of an automotive engine. In N. Lynch, & B. Krogh (Eds.), *Lecture notes in computer science*: *Vol. 1790. Hybrid systems: Computation and control* (pp. 32–44). Springer Verlag.

Belta, C., Finin, P., Habets, L., Halász, A., Imielinski, M., Kumar, V., et al. (2004). Understanding the bacterial stringent response using reachability analysis of hybrid systems. In R. Alur, & G. J. Pappas (Eds.), *Lecture notes in computer science*: *Vol. 2993. Hybrid systems: Computation and control* (pp. 111–126). Springer Verlag.

Bertsekas, D. P. (1972). Infinite-time reachability of state-space regions using feedback control. *IEEE Transactions on Automatic Control*, *AC-17*(5), 604–613.

Bertsekas, D. P., & Shreve, S. E. (1996). *Stochastic optimal control: The discrete-time case*. Athena Scientific.

Bujorianu, M. L. (2004). Extended stochastic hybrid systems and their reachability problem. In R. Alur, & G. Pappas (Eds.), *Lecture notes in computer science*: *Vol. 2993. Hybrid systems: Computation and control* (pp. 234–249). Springer Verlag.

Bujorianu, M. L., & Lygeros, J. (2004). General stochastic hybrid systems: Modelling and optimal control. In *Proceedings of the 43rd IEEE conference on decision and control*.

Davis, M. H. A. (1993). *Markov models and optimization*. London: Chapman & Hall/CRC Press.

Digailova, I. A., & Kurzhanski, A. B. (2005). Reachability analysis under control-dependent stochastic noise. In *Proceedings of the 16th IFAC world congress*.

Fehnker, A., & Ivančić, F. (2004). Benchmarks for hybrid systems verifications. In R. Alur, & G. J. Pappas (Eds.), *Lecture notes in computer science*: *Vol. 2993. Hybrid systems: Computation and control* (pp. 326–341). Springer Verlag.

Ghosh, M. K., Araposthasis, A., & Marcus, S. I. (1997). Ergodic control of switching diffusions. *SIAM Journal of Control and Optimization*, *35*(6), 1952–1988.

Girard, A. (2005). Reachability of uncertain linear systems using zonotopes. In M. Morari, & L. Thiele (Eds.), *Lecture notes in computer science*: *Vol. 3414. Hybrid systems: Computation and control* (pp. 291–305). Springer Verlag.

Girard, A., Julius, A., & Pappas, G. J. (2006). Approximate simulation relations for hybrid systems. In *Proceedings of the 2nd IFAC conference on analysis and design of hybrid systems*.

Hu, J., Lygeros, J., & Sastry, S. (2000). Towards a theory of stochastic hybrid systems. In N. Lynch, & B. Krogh (Eds.), *Lecture notes in computer science*: *Vol. 1790. Hybrid systems: Computation and control* (pp. 160–173). Springer Verlag.

Hu, J., Prandini, M., & Sastry, S. (2005). Aircraft conflict prediction in the presence of a spatially correlated wind field. *IEEE Transactions on Intelligent Transportation Systems*, *3*, 326–340.

Katoen, J. P. (2006). Stochastic model checking. In C. G. Cassandras, & J. Lygeros (Eds.), *Automation and control engineering series*: *Vol. 24. Stochastic hybrid systems* (pp. 79–106). Taylor & Francis Group/CRC Press.

Kumar, P. R., & Varaiya, P. P. (1986). *Stochastic systems: Estimation, identification, and adaptive control*. New Jersey: Prentice Hall, Inc.

Kurzhanski, A. B., & Varaiya, P. (2002). On reachability under uncertainty. *SIAM Journal of Control and Optimization*, *41*(1), 181–216.

Kushner, H. J., & Dupuis, P. G. (2001). *Numerical methods for stochastic control problems in continuous time*. New York: Springer-Verlag.

Lygeros, J. (2004). On reachability and minimum cost optimal control. *Automatica*, *40-6*, 317–927.

Lygeros, J., Tomlin, C., & Sastry, S. (1999). Controllers for reachability specifications for hybrid systems. *Automatica*, *35*(3), 349–370.

Lygeros, J., & Watkins, O. (2003). Stochastic reachability for discrete time systems: an application to aircraft collision avoidance. In *Proceedings of the 42nd IEEE conference of decision and control*.

Malhame, R., & Chong, C.-Y. (1985). Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system. *IEEE Transactions on Automatic Control*, *30*(9), 854–860.

Mitchell, I., & Templeton, J. (2005). A toolbox of Hamilton-Jacobi solvers for analysis of nondeterministic continuous and hybrid systems. In M. Morari, & L. Thiele (Eds.), *LNCIS*: *Vol. 3414. Hybrid systems: Computation and control* (pp. 480–494). Springer Verlag.

Mitchell, I., & Tomlin, C. (2000). Level set methods for computation in hybrid systems. In B. Krogh, & N. Lynch (Eds.), *Lecture notes in computer science*, *Hybrid systems: Computation and control* (pp. 310–323). Springer Verlag.

Picasso, B., & Bicchi, A. (2005). Control synthesis for practical stabilization of quantized linear systems. *Rendiconti Seminario Matematico Università di Torino*, *63*(4), 397–410.

Prajna, S., Jadbabaie, A., & Pappas, G. J. (2007). A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Transactions on Automatic Control*, *52*(8), 1415–1429.

Prandini, M., & Hu, J. (2006a). A stochastic approximation method for reachability computations. In H. A. P. Blom, & J. Lygeros (Eds.), *Lecture notes in control and informations sciences*: *Vol. 337*. *Stochastic hybrid systems: Theory and safety applications* (pp. 107–139). Springer.

Prandini, M., & Hu, J. (2006b). Stochastic reachability: Theory and numerical approximation. In C. G. Cassandras, & J. Lygeros (Eds.), *Automation and control engineering series*: *Vol. 24. Stochastic hybrid systems* (pp. 107–138). Taylor & Francis Group/CRC Press.

Prandini, M., Hu, J., Lygeros, J., & Sastry, S. (2000). A probabilistic approach to aircraft conflict detection. *IEEE Transactions on Intelligent Transportation Systems*, *1*(4), 199–220.

Puterman, M. L. (1994). *Markov decision processes*. John Wiley & Sons, Inc.

Stursberg, O., & Krogh, B. H. (2003). Effcient representation and computation of reachable sets for hybrid systems. In A. Pnueli, & O. Maler (Eds.), *Lecture notes in computer science*: *Vol. 2623. Hybrid systems: Computation and control* (pp. 482–497). Springer Verlag.

Tomlin, C., Lygeros, J., & Sastry, S. (1998). Synthesizing controllers for nonlinear hybrid systems. In T. Henzinger, & S. Sastry (Eds.), *Lecture notes in computer science*: *Vol. 1386*. *Hybrid systems: Computation and control* (pp. 360–373). Springer Verlag.

Tomlin, C., Mitchell, I., Bayen, A., & Oishi, M. (2003). Computational techniques for the verification and control of hybrid systems. *Proceedings of the IEEE*, *91*(7), 986–1001.

Alessandro Abate received the Laurea degree in Electrical Engineering from the University of Padova in 2002, and the M.S. and Ph.D. degrees in Electrical Engineering and Computer Sciences from the University of California, Berkeley, in 2004 and 2007 respectively. He is currently a Postdoctoral Researcher at the Department of Aeronautics and Astronautics at Stanford University. His research interests are in the analysis, control, and verification of probabilistic and hybrid systems, and their application in systems biology.

Maria Prandini received the Laurea degree in electrical engineering from the Politecnico of Milano, Italy, in 1994, and the Ph.D. degree in electrical engineering from the University of Brescia, Italy, in 1998. She was a visiting postdoctoral researcher at the University of California at Berkeley, from 1998 to 2000. Since December 2002 she is Assistant Professor at the Dipartimento di Elettronica e Informazione of the Politecnico di Milano. Her research interests include adaptive control of stochastic systems, multi-agent coordination and control, air traffic management, and stochastic hybrid systems. She is currently discussion editor of the European Journal of Control, and member of the IFAC Technical Committee on Stochastic Systems and of the IEEE Control Systems Society Conference Editorial Board.

John Lygeros received a B.Eng. degree in electrical engineering and an M.Sc. degree in automatic control from Imperial College, London, U.K., in 1990 and 1991, respectively. In 1996, he completed a Ph.D. degree at the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley. He subsequently held a series of Postdoctoral research appointments with the National Automated Highway Systems Consortium, Massachusetts Institute of Technology, and the University of California, Berkeley. In parallel, he was also a part-time Research Engineer with SRI International, Menlo Park, CA, and a Visiting Professor with the Department of Mathematics, Université de Bretagne Occidentale, Brest, France. Between July 2000 and March 2003, he was a University Lecturer with the Department of Engineering, University of Cambridge and a Fellow of Churchill College, Cambridge, U.K.. Between March 2003 and July 2006, he was an Assistant Professor with the Department of Electrical and Computer Engineering, University of Patras, Patras, Greece. In July 2006, he joined ETH Zurich, Switzerland, as an Associate Professor with the Automatic Control Laboratory. His research interests include modeling, analysis, and control of hierarchical hybrid systems with applications to biochemical networks and large-scale engineering systems such as automated highways and air traffic management.

Shankar Sastry received a B.Tech. from the Indian Institute of Technology, Bombay, 1977, a M.S. in EECS, M.A. in Mathematics and Ph.D. in EECS from UC Berkeley, 1979, 1980, and 1981 respectively. S. Shankar Sastry is currently dean of the College of Engineering. He was formerly the Director of CITRIS (Center for Information Technology Research in the Interest of Society) and the Banatao Institute @ CITRIS Berkeley. He served as chair of the EECS department from January, 2001 through June 2004. In 2000, he served as Director of the Information Technology Office at DARPA. From 1996-1999, he was the Director of the Electronics Research Laboratory at Berkeley, an organized research unit on the Berkeley campus conducting research in computer sciences and all aspects of electrical engineering. He is the NEC Distinguished Professor of Electrical Engineering and Computer Sciences and holds faculty appointments in the Departments of Bioengineering, EECS and Mechanical Engineering. Prior to joining the EECS faculty in 1983 he was a professor at MIT.