Automatica 49 (2013) 2665–2674

Contents lists available at ScienceDirect

Automatica

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

A stochastic games framework for verification and control of discrete time stochastic hybrid systems☆

Jerry Ding [a,1,2], Maryam Kamgarpour [b,2], Sean Summers [b], Alessandro Abate [c], John Lygeros [b], Claire Tomlin [a]

[a] Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720, USA

[b] Automatic Control Laboratory, Department of Information Technology and Electrical Engineering, Swiss Federal Institute of Technology (ETH) Zürich, CH-8092 Zürich, Switzerland

[c] Department of Computer Science, University of Oxford, Oxford, OX1 3QD, UK

ARTICLE INFO

Article history:
Received 25 May 2012
Received in revised form
28 December 2012
Accepted 21 May 2013
Available online 20 June 2013

Keywords:
Hybrid systems
Stochastic systems
Dynamic games
Controller synthesis
Reachability

ABSTRACT

We describe a framework for analyzing probabilistic reachability and safety problems for discrete time stochastic hybrid systems within a dynamic games setting. In particular, we consider finite horizon zero-sum stochastic games in which a control has the objective of reaching a target set while avoiding an unsafe set in the hybrid state space, and a rational adversary has the opposing objective. We derive an algorithm for computing the maximal probability of achieving the control objective, subject to the worst-case adversary behavior. From this algorithm, sufficient conditions of optimality are also derived for the synthesis of optimal control policies and worst-case disturbance strategies. These results are then specialized to the safety problem, in which the control objective is to remain within a safe set. We illustrate our modeling framework and computational approach using both a tutorial example with jump Markov dynamics and a practical application in the domain of air traffic management.

1. Introduction

In application scenarios ranging from air traffic management (Sastry et al., 1995; Tomlin, Pappas, & Sastry, 2002), automotive control (Balluchi, Benvenuti, Di Benedetto, Pinello, & Sangiovanni-Vincentelli, 2000), systems biology (Ghosh & Tomlin, 2004; Lincoln & Tiwari, 2004), to bipedal walking (Ames, Sinnet, & Wendel, 2009), the behavior of the system can be described in terms of a hybrid system abstraction in which the system state evolves both in the discrete and continuous domain. While the discrete state can be used to capture the qualitative behavior of the system, for example the operating modes of a flight management system or the phases of a walking cycle, the continuous state can be used to capture quantitative characteristics such as the velocity and heading of the aircraft or the joint angles of a biped. When the evolution of the discrete and continuous state can be modeled probabilistically, for example through analysis of statistical data, then a natural modeling framework is that of a stochastic hybrid system (SHS) (Glover & Lygeros, 2004; Hespanha, 2004; Hu, Lygeros, & Sastry, 2000).

For a controlled SHS, the performance of the closed-loop system can be measured in terms of the probability that the system trajectory obeys desired specifications. Of interest to safety-critical applications are probabilistic reachability and safety problems in which the control objective is to maximize the probability of reaching a desired target set or of remaining within a given safe set. In the continuous-time case, a theoretical upper bound on reachable event probabilities is derived in Bujorianu (2004) using Dirichlet forms. The temporal evolution of the probability density function of the hybrid state has been characterized through generalized Fokker–Planck equations (Bect, Phulpin, Baili, & Fleury, 2006). Optimal control of SHS is considered in Bensoussan and Menaldi (2000) and quasi-variational inequalities based on dynamic programming are derived for the optimal trajectory. An optimal control approach towards reachability analysis is discussed in Koutsoukos and Riley (2006) and Mohajerin Esfahani, Chatterjee, and Lygeros (2011), in which the solutions of probabilistic reachability and safety problems are derived in terms of the viscosity solutions of appropriate Hamilton–Jacobi–Bellman equations. To address issues of computational complexity, the authors in Hu,

☆ The material in this paper was partially presented at the 50th IEEE Conference on Decision and Control (CDC), December 12–15, 2011, Orlando, Florida, USA. This paper was recommended for publication in revised form by Associate Editor Jan Komenda under the direction of Editor Ian R. Petersen.

E-mail addresses: jding@eecs.berkeley.edu (J. Ding), mkamgar@control.ee.ethz.ch (M. Kamgarpour), summers@control.ee.ethz.ch (S. Summers), Alessandro.Abate@cs.ox.ac.uk (A. Abate), lygeros@control.ee.ethz.ch (J. Lygeros), tomlin@eecs.berkeley.edu (C. Tomlin).

[1] Tel.: +1 414 630 4868; fax: +1 510 642 2718.

[2] The first two authors contributed equally.

0005-1098/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.automatica.2013.05.025

Prandini, and Sastry (2005) propose a Markov chain approximation of SHS using methods from Kushner and Dupuis (1992), while in Prajna, Jadbabaie, and Pappas (2007), the authors discuss an approach for computing an upper bound on the safety probability using barrier certificates. For discrete-time stochastic hybrid systems (DTSHS), a theoretical framework for the study of probabilistic safety problems is established in Abate, Prandini, Lygeros, and Sastry (2008). These results are generalized in Summers and Lygeros (2010) to address the reach–avoid problem, in which the control objective is to reach a desired target set, while remaining within a safe set. Considerations for time-varying and stochastic sets are discussed in Abate, Amin, Prandini, Lygeros, and Sastry (2006) and Summers, Kamgarpour, Lygeros, and Tomlin (2011), respectively.

While much of the previous work has studied optimal control formulations of probabilistic reachability and safety problems, in which the evolution of the system is only subject to inputs by the control, we consider in this paper an extension to the case of zero-sum stochastic games, in which the system dynamics are also subject to inputs by an adversary, whose objectives are opposed to that of the control. In particular, generalizing the results in Abate et al. (2008) and Summers and Lygeros (2010) for optimal control of DTSHS, our recent work in Kamgarpour et al. (2011) introduced a framework for the study of max–min probabilistic reachability problems within the context of a stochastic game model of DTSHS. This is motivated by practical applications such as conflict resolution in air traffic management (Tomlin et al., 2002) and secure control of networked systems subject to external attacks (Amin, Cardenas, & Sastry, 2009), in which the intent of certain rational agents may be uncertain. In addition, the framework is applicable to robust control applications, in which unmodeled dynamics or bounded disturbances are to be accounted for in a worst-case fashion.

In this article, we expand upon our work in Kamgarpour et al. (2011) by providing a thorough exposition of the theoretical results, along with a detailed analysis of several examples. Most importantly, we present a detailed proof for the main theorem in Kamgarpour et al. (2011), which provides a dynamic programming approach for computing the maximal probability of satisfying a reach–avoid specification, subject to the worst-case adversary behavior. This proof also allows us to derive sufficient conditions of optimality for the synthesis of optimal policies for the control and the adversary. Furthermore, we demonstrate how these results can be specialized to address the safety problem, by computing the minimal probability that the system state reaches an unsafe subset of the state space. Finally, we also provide detailed discussions of both a tutorial analytical example as well as a practical numerical example in order to illustrate the application of the proposed methodology.

Our main contribution is a theoretic framework for the study of probabilistic reachability and safety problems for DTSHS within the setting of zero-sum stochastic games, extending previous work on the optimal control framework in Abate et al. (2008) and Summers and Lygeros (2010). It is important to note that such an extension requires addressing several subtle and yet challenging issues that are unique to stochastic games: (1) the choice of an appropriate information pattern; (2) the measurability of value functions under max–min operations; (3) the existence of equilibrium strategies within appropriate classes.

We will now briefly elaborate on these issues. First, depending on what information one assumes is exchanged between the control and the adversary in a zero-sum game, one can arrive at drastically different problem formulations, with correspondingly different game values and interpretations of solution strategies. Motivated by an interest in robust control, this work considers an asymmetric information pattern which favors the adversary, leading to a max–min (Gonzalez-Trejo, Hernandez-Lerma, & Hoyos-Reyes, 2002) or Stackelberg game formulation (Breton, Alj, & Haurie, 1988) of the zero-sum game. Second, measurability of value functions, which are vital for ensuring that the probabilities of interest can be computed recursively by a dynamic programming procedure, is often significantly more difficult to establish in a stochastic game setting as compared with an optimal control setting, due to nested maximization and minimization (Nowak, 1985). Thus, formal proofs of dynamic programming results require analysis tools and proof techniques stemming from the field of non-cooperative stochastic games (Gonzalez-Trejo et al., 2002; Kumar & Shiau, 1981; Maitra & Sudderth, 1998; Nowak, 1985; Rieder, 1991). Finally, it is a well-known fact that equilibrium strategies for zero-sum stochastic games need not exist within the space of pure (i.e. deterministic) strategies (Maitra & Parthasarathy, 1970; Shapley, 1953). On the other hand, by assuming an asymmetric information pattern, as well as continuity and compactness properties on the system model, we show that there exists a solution to the max–min reachability problem within the space of pure strategies, albeit at the cost of conservativeness.

The article is organized as follows. In Section 2, we discuss the model for a discrete-time stochastic hybrid game (DTSHG). In Section 3, we give formal stochastic game formulations of the probabilistic reach–avoid and safety problems. In Section 4, we state and prove our main result for computing the max–min reach–avoid probability, and give sufficient conditions of optimality for both the control and the adversary. This is followed by the specialization of this result to the safety problem. Throughout, we illustrate the terminology and methodology using a simple jump Markov system. In Section 5, we apply our framework to a practical example in air traffic management. Finally, concluding remarks along with directions for future work are given in Section 6.

2. Discrete-time stochastic hybrid game

The model for a discrete-time stochastic hybrid game (DTSHG) proposed here is an extension of the discrete-time stochastic hybrid systems (DTSHS) model proposed in Abate et al. (2008) and Summers and Lygeros (2010) to a two-player stochastic game setting. As in previous work, we require the stochastic transition kernels to be Borel-measurable and denote by $\mathcal{B}(\cdot)$ the Borel σ-algebra. This condition ensures that the probabilities of interest can be computed by integration of the transition kernels over a hybrid state space. Following standard conventions, we refer to the control as Player I and the adversary as Player II.

Definition 1 (*DTSHG*)**.** A discrete-time stochastic hybrid game between two players is a tuple $\mathcal{H} = (\mathcal{Q}, n, \mathcal{A}, \mathcal{D}, \tau_v, \tau_q, \tau_r)$, defined as follows.

- *Discrete state space* $\mathcal{Q} := \{q_1, q_2, \ldots, q_m\}$, $m \in \mathbb{N}$;
- *Dimension of continuous state space* $n : \mathcal{Q} \to \mathbb{N}$: a map which assigns to each discrete state $q \in \mathcal{Q}$ the dimension of the continuous state space. The hybrid state space is given by $X := \bigcup_{q\in\mathcal{Q}} \{q\} \times \mathbb{R}^{n(q)}$;
- *Player* I *controls* $\mathcal{A}$: a nonempty, compact Borel space;
- *Player* II *controls* $\mathcal{D}$: a nonempty, compact Borel space;
- *Continuous state transition kernel* $\tau_v(dv'|(q, v), a, d)$: a collection of Borel-measurable stochastic kernels on $\mathbb{R}^{n(q)}$ given $\mathbb{R}^{n(q)} \times \mathcal{A} \times \mathcal{D}$, which assigns to each $x = (q, v) \in X$, $a \in \mathcal{A}$, and $d \in \mathcal{D}$ a probability measure on the Borel space $(\mathbb{R}^{n(q)}, \mathcal{B}(\mathbb{R}^{n(q)}))$;
- *Discrete state transition kernel* $\tau_q(q'|x, a, d)$: a Borel-measurable stochastic kernel on $\mathcal{Q}$ given $X \times \mathcal{A} \times \mathcal{D}$, which assigns to each $x \in X$, $a \in \mathcal{A}$, and $d \in \mathcal{D}$ a probability distribution over $\mathcal{Q}$;

- *Reset transition kernel* $\tau_r(dv'|(q,v),a,d,q')$: a collection of Borel-measurable stochastic kernels on $\mathbb{R}^{n(q')}$ given $\mathbb{R}^{n(q)} \times \mathcal{A} \times \mathcal{D}$, which assigns to each $x = (q,v) \in X$, $a \in \mathcal{A}$, $d \in \mathcal{D}$, and $q' \in \mathcal{Q} \setminus \{q\}$ a probability measure on the Borel space $(\mathbb{R}^{n(q')}, \mathcal{B}(\mathbb{R}^{n(q')}))$.

In contrast with the single-player case, the stochastic transition kernels in a DTSHG are affected by the inputs of two agents with possibly differing objectives. In particular, we assume that Player I and Player II are non-cooperative and consider a conservative decision model in which the actions of Player II may be chosen in a rational fashion based upon the actions of Player I.

Definition 2. A Markov policy for player I is a sequence $\mu = (\mu_0, \mu_1, \ldots, \mu_{N-1})$ of Borel-measurable maps $\mu_k : X \to \mathcal{A}$, $k = 0, 1, \ldots, N-1$. The set of all admissible Markov policies for Player I is denoted by $\mathcal{M}_a$.

Definition 3. A Markov strategy for Player II is a sequence $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_{N-1})$ of Borel-measurable maps $\gamma_k : X \times \mathcal{A} \to \mathcal{D}$, $k = 0, 1, \ldots, N-1$. The set of all admissible Markov strategies for Player II is denoted by Γ_d.

The scenario described here is a common setting in robust control problems in which the control selects inputs in anticipation of the worst-case response by an adversary or a disturbance. More formally, this can be interpreted as a zero-sum Stackelberg game in which Player I is the leader. Due to the asymmetry in information in a Stackelberg game, equilibrium strategies of a zero-sum game can be typically chosen to be deterministic rather than randomized (Breton et al., 1988). We note, however, that in a zero-sum stochastic game with symmetric information (the actions of Player I are not revealed to Player II), the existence of a non-cooperative equilibrium in general requires randomized strategies (see for example Maitra & Parthasarathy, 1970; Shapley, 1953). Furthermore, if one were to consider transition probabilities and general utility functions which depend on the entire history of the game, it may also be necessary to broaden the class of player strategies to encompass non-Markov policies (Maitra & Sudderth, 1998; Rieder, 1991). However, as shown in Rieder (1991), when the transition probabilities are Markovian and the utility function is sum-multiplicative (as in our case), it is sufficient to consider the class of Markov control policies.

For a given initial condition $x(0) = (q_0, v_0) \in X$, Player I policy $\mu \in \mathcal{M}_a$, and Player II strategy $\gamma \in \Gamma_d$, the semantics of a DTSHG can be described as follows. At time step k, each player obtains a measurement of the current system state $x(k) = (q(k), v(k)) \in X$. Using this information, Player I selects a control input $a(k) = \mu_k(x(k))$, following which Player II selects a disturbance input $d(k) = \gamma_k(x(k), a(k))$. The discrete state is then updated according to the discrete transition kernel as $q(k+1) \sim \tau_q(\cdot|x(k), a(k), d(k))$. If the discrete state remains the same, namely $q(k+1) = q(k)$, then the continuous state is updated according to the continuous state transition kernel as $v(k+1) \sim \tau_v(\cdot|x(k), a(k), d(k))$. On the other hand, if there is a discrete jump, the continuous state is instead updated according to the reset transition kernel as $v(k+1) \sim \tau_r(\cdot|x(k), a(k), d(k), q(k+1))$.

Following this description, we can compose the transition kernels τ_v, τ_q, and τ_r to form a Borel-measurable hybrid state transition kernel $\tau_q(dx'|x,a,d)$ which describes the evolution of the hybrid state under the influence of Player I and Player II inputs:

$$\tau\big((q', dv')|(q,v),a,d\big) := \begin{cases} \tau_q(q|(q,v),a,d)\tau_v(dv'|(q,v),a,d), & \text{if } q' = q \\ \tau_q(q'|(q,v),a,d)\tau_r(dv'|(q,v),a,d,q'), & \text{if } q' \neq q. \end{cases} \quad (1)$$

Using the transition kernel τ, we can now give a formal definition for the executions of a DTSHG.

Definition 4. Let $\mathcal{H}$ be a DTSHG and $N \in \mathbb{N}$ be a finite time horizon. For a given $\mu \in \mathcal{M}_a$, $\gamma \in \Gamma_d$, and $x_0 \in X$, a stochastic process $\{x(k), k = 0, \ldots, N\}$ with values in X is an execution of $\mathcal{H}$ if its sample paths are generated according to Algorithm 1.

Algorithm 1 DTSHG Execution

Require: Initial hybrid state $x_0 \in X$, Player I policy $\mu \in \mathcal{M}_a$, Player II strategy $\gamma \in \Gamma_d$;
Set $x(0) = x_0$;
for $k = 0$ to $N-1$ **do**
Set $a(k) = \mu_k(x(k))$;
Set $d(k) = \gamma_k(x(k), a(k))$;
Extract from X a value x_{k+1} for $x(k+1)$ according to $\tau(\cdot|x(k), a(k), d(k))$;
end for
return Sample Path $\{x_k, \ k = 0, \ldots, N\}$.

By this definition, the execution of a DTSHG is a time inhomogeneous stochastic process on the sample space $\Omega = X^{N+1}$, endowed with the canonical product topology $\mathcal{B}(\Omega) := \prod_{k=1}^{N+1} \mathcal{B}(X)$. The evolution of the closed-loop hybrid state trajectory can be described in terms of the transition kernels $\tau^{\mu_k,\gamma_k}(\cdot|x) := \tau(\cdot|x, \mu_k(x), \gamma_k(x, \mu_k(x)))$, $k = 0, \ldots, N$. By Proposition 7.28 of Bertsekas and Shreve (1978), for a given $x_0 \in X$, $\mu \in \mathcal{M}_a$, $\gamma \in \Gamma_d$, these stochastic kernels induce a unique probability measure $P_{x_0}^{\mu,\gamma}$ on Ω as defined by

$$P_{x_0}^{\mu,\gamma}(X_0 \times X_1 \times \cdots \times X_N) := \int_{X_0} \int_{X_1} \cdots \int_{X_N} \prod_{k=0}^{N-1} \tau^{\mu_k,\gamma_k}(dx'_{k+1}|x'_k)\delta_{x_0}(dx'_0), \quad (2)$$

where $X_0, X_1, \ldots, X_N \in \mathcal{B}(X)$ are Borel sets and δ_{x_0} denotes the probability measure on X which assigns unit mass to the point $x_0 \in X$.

2.1. Example—2-mode jump Markov system

Consider a simple jump Markov system with two modes of operation $\mathcal{Q} = \{q_1, q_2\}$. The transitions between the discrete modes are modeled probabilistically, with the probability of dwelling in mode q_i given by p_i, $i = 1, 2$. While in mode q_i, a continuous state $v \in \mathbb{R}$ evolves according to a stochastic difference equation $v(k+1) = f_i(v(k), a(k), d(k), \eta(k))$, defined as follows:

$$f_i(v(k), a(k), d(k), \eta(k)) := \begin{cases} 2v(k) + a(k) + d(k) + \eta(k), & \text{if } i = 1 \\ \frac{1}{2}v(k) + a(k) + d(k) + \eta(k), & \text{if } i = 2. \end{cases} \quad (3)$$

where a and d are Player I and Player II inputs, and η is a random variable. It is assumed that the players have identical capabilities, with $a, d \in [-1, 1]$. The noise is modeled by a uniform distribution $\eta \sim \mathcal{U}[-1, +1]$.

Under the DTSHG modeling framework, the hybrid state space is $X = \{q_1, q_2\} \times \mathbb{R}$, and the players' input spaces are $\mathcal{A} = \mathcal{D} = [-1, 1]$. The discrete transition kernel τ_q is derived as $\tau_q(q_1|(q_1, v), a, d) = p_1$, $\tau_q(q_2|(q_1, v), a, d) = 1 - p_1$, $\tau_q(q_1|(q_2, v), a, d) = 1 - p_2$, $\tau_q(q_2|(q_2, v), a, d) = p_2$. The continuous transition kernel τ_v can be derived from the continuous state dynamics (3) as $\tau_v(dv'|(q_1, v), a, d) \sim \mathcal{U}[2v + a + d - 1, 2v + a + d + 1]$, $\tau_v(dv'|(q_2, v), a, d) \sim \mathcal{U}[\frac{1}{2}v + a + d - 1, \frac{1}{2}v + a + d + 1]$. Finally, the reset transition kernel is given by $\tau_r(dv'|(q,v), a, d, q') = \tau_v(dv'|(q,v), a, d)$.

3. Probabilistic reach–avoid and safety problems for DTSHG

Within the context of a DTSHG model, we discuss in this section stochastic game formulations of the probabilistic reach–avoid and safety problems, as introduced by Summers and Lygeros (2010) and Abate et al. (2008), respectively, in an optimal control setting. First, we consider a reach–avoid problem in which the objective of Player I (the control) is to steer the hybrid system state into a desired target set, while avoiding a set of unsafe states, and the objective of Player II (the adversary) is to prevent Player I from doing so.

More precisely, suppose that a Borel set $K \in \mathcal{B}(X)$ is given as the target set, while $K' \in \mathcal{B}(X)$ is given as the safe set, with $K \subseteq K'$. Then the probability that the state trajectory $(x_0, x_1, \ldots, x_N)$ reaches K while staying within K' under fixed $\mu \in \mathcal{M}_a$ and $\gamma \in \Gamma_d$ is given by

$$\begin{aligned} r^{\mu,\gamma}_{x_0}(K, K') &:= P^{\mu,\gamma}_{x_0}(\{(x_0, \ldots, x_N) : \exists j \in \{0, 1, \ldots, N\}, \\ &\qquad (x_j \in K) \wedge (x_i \in K',\ \forall i \in \{0, 1, \ldots, j\})\}) \\ &= \sum_{j=0}^{N} P^{\mu,\gamma}_{x_0}((K' \setminus K)^j \times K \times X^{N-j}). \end{aligned} \tag{4}$$

Following a similar procedure as in Summers and Lygeros (2010), this probability can be rewritten as

$$r^{\mu,\gamma}_{x_0}(K, K') = E^{\mu,\gamma}_{x_0}\left[\mathbf{1}_K(x_0) + \sum_{j=1}^{N}\left(\prod_{i=0}^{j-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_j)\right], \tag{5}$$

where $E^{\mu,\gamma}_{x_0}$ denotes the expectation with respect to the probability measure $P^{\mu,\gamma}_{x_0}$, and $\mathbf{1}_{X'}$ denotes the indicator function of a set $X' \subseteq X$. Now define the worst-case reach–avoid probability under a Player I policy μ as

$$r^{\mu}_{x_0}(K, K') := \inf_{\gamma \in \Gamma_d} r^{\mu,\gamma}_{x_0}(K, K'). \tag{6}$$

The reach–avoid problem for a DTSHG is as follows.

Problem 1. Given a DTSHG $\mathcal{H}$, target set $K \in \mathcal{B}(X)$, and safe set $K' \in \mathcal{B}(X)$ such that $K \subseteq K'$:

(I) Compute the max–min reach–avoid probability $r^*_{x_0}(K, K') := \sup_{\mu\in\mathcal{M}_a} r^{\mu}_{x_0}(K, K'), \forall x_0 \in X$;
(II) Find a max–min control policy $\mu^* \in \mathcal{M}_a$, whenever it exists, such that $r^*_{x_0}(K, K') = r^{\mu^*}_{x_0}(K, K'), \forall x_0 \in X$.
(III) Find a worst-case adversary strategy $\gamma^* \in \mathcal{M}_a$, whenever it exists, such that $r^*_{x_0}(K, K') = r^{\mu^*,\gamma^*}_{x_0}(K, K'), \forall x_0 \in X$.

We now consider a safety problem in which the objective of Player I (the control) is to keep the system state within a safe set, and the objective of Player II (the adversary) is to steer the system state into the unsafe set. Following a similar approach as in Summers and Lygeros (2010), one can formulate the safety problem as a special case of the reach–avoid problem. This stems from the observation that the hybrid state x_k remains within a safe set S *for all* k if and only if it does not reach the unsafe set $X \setminus S$ *for any* k. Mathematically speaking, for fixed $\mu \in \mathcal{M}_a$ and $\gamma \in \Gamma_d$, the safety probability is given by

$$\begin{aligned} p^{\mu,\gamma}_{x_0}(S) &:= P^{\mu,\gamma}_{x_0}(\{(x_0, \ldots, x_N) : x_k \in S,\ \forall k \in \{0, 1, \ldots, N\}\}) \\ &= 1 - r^{\mu,\gamma}_{x_0}(X \setminus S, X). \end{aligned} \tag{7}$$

The safety problem for a DTSHG is then characterized by the following max–min value function.

$$p^*_{x_0}(S) := \sup_{\mu\in\mathcal{M}_a}\inf_{\gamma\in\Gamma_d} p^{\mu,\gamma}_{x_0}(S). \tag{8}$$

Similarly as in the single-player case, both the reach–avoid and safety problems can readily be modified to account for time-varying (Abate et al., 2006) and stochastic (Summers et al., 2011) target sets and safe sets. For simplicity of notation, we will focus here on static and deterministic sets.

4. Dynamic programming solution

4.1. Main theorem

For our theoretical derivations, we impose the following regularity assumptions on the stochastic kernels.

Assumption 1. (a) For each $x = (q, v) \in X$ and $E_1 \in \mathcal{B}(\mathbb{R}^{n(q)})$, the function $(a, d) \to \tau_v(E_1|x, a, d)$ is continuous;
(b) For each $x = (q, v) \in X$ and $q' \in \mathcal{Q}$, the function $(a, d) \to \tau_q(q'|x, a, d)$ is continuous;
(c) For each $x = (q, v) \in X$, $q' \in \mathcal{Q} \setminus \{q\}$, and $E_2 \in \mathcal{B}(\mathbb{R}^{n(q')})$, the function $(a, d) \to \tau_r(E_2|x, a, d, q')$ is continuous.

The need for continuity assumptions on stochastic kernels commonly arise in the stochastic games literature (see for example Gonzalez-Trejo et al., 2002; Kumar & Shiau, 1981; Maitra & Parthasarathy, 1970; Nowak, 1985), due to the difficulties in ensuring the measurability of value functions under max–min dynamic programming operations. Following the approach in Nowak (1985) and Rieder (1991), we only assume continuity of the stochastic kernels in the actions of Player I and Player II, but not necessarily in the system state. This allows for stochastic hybrid systems in which transition probabilities change abruptly with changes in the system state. Furthermore, if the action spaces $\mathcal{A}$ and $\mathcal{D}$ are finite or countable, then the assumptions are satisfied under the discrete topology on $\mathcal{A}$ and $\mathcal{D}$. Also, the assumptions on τ_v and τ_r are satisfied if these kernels admit density functions that are continuous in the player inputs.

In order to provide a solution to Problem 1, we define a max–min dynamic programming operator T which takes as its argument a Borel-measurable function $J : X \to [0, 1]$ and produces another real-valued function on X:

$$T(J)(x) := \mathbf{1}_K(x) + \sup_{a\in\mathcal{A}}\inf_{d\in\mathcal{D}} \mathbf{1}_{K'\setminus K}(x)H(x, a, d, J),$$

$$\text{where } H(x, a, d, J) := \int_X J(y)\tau(dy|x, a, d). \tag{9}$$

Theorem 1. *Let $\mathcal{H}$ be a DTSHG satisfying Assumption 1. Let $K, K' \in \mathcal{B}(X)$ be Borel sets such that $K \subseteq K'$. Let the operator T be defined as in (9). Then the composition $T^N = T \circ T \circ \cdots \circ T$ (N times) is well-defined and*

(a) $r^*_{x_0}(K, K') = T^N(\mathbf{1}_K)(x_0), \forall x_0 \in X$;
(b) *There exists a Player I policy $\mu^* \in \mathcal{M}_a$ and a Player II strategy $\gamma^* \in \Gamma_d$ satisfying*

$$r^{\mu,\gamma^*}_{x_0}(K, K') \le r^*_{x_0}(K, K') \le r^{\mu^*,\gamma}_{x_0}(K, K'), \tag{10}$$

$\forall x_0 \in X$, $\mu \in \mathcal{M}_a$, and $\gamma \in \Gamma_d$. In particular, μ^* is a max–min control policy, and γ^* is a worst-case adversary strategy.
(c) *Let $V^*_N = \mathbf{1}_K$, $V^*_k = T^{N-k}(\mathbf{1}_K)$, $k = 0, 1, \ldots, N-1$. If $\mu^* \in \mathcal{M}_a$ is a Player I policy which satisfies*

$$\mu^*_k(x) \in \arg\max_{a\in\mathcal{A}}\inf_{d\in\mathcal{D}} H(x, a, d, V^*_{k+1}), \tag{11}$$

$\forall x \in K' \setminus K$, $k = 0, 1, \ldots, N-1$, then μ^* is a max–min control policy. If $\gamma^* = (\gamma^*_0, \gamma^*_1, \ldots, \gamma^*_{N-1}) \in \Gamma_d$ is a Player II strategy

which satisfies

$$\gamma_k^*(x,a) \in \arg\min_{d\in\mathcal{D}} H(x,a,d,V_{k+1}^*), \tag{12}$$

$\forall x \in K'\backslash K,\ a\in\mathcal{A},\ k=0,1,\ldots,N-1$, *then* γ^* *is a worst-case adversary strategy.*

First, we will present a recursive procedure for computing the reach–avoid probability $r_{x_0}^{\mu,\gamma}(K,K')$, under fixed choices of Player I policy $\mu\in\mathcal{M}_a$ and Player II strategy $\gamma\in\Gamma_d$. Consider the cost-to-go functions $V_k^{\mu,\gamma}: X\to[0,1]$, $k=0,\ldots,N$, defined as

$$V_N^{\mu,\gamma}(x_N) := \mathbf{1}_K(x_N), \tag{13}$$

$$V_k^{\mu,\gamma}(x_k) := E_{x_k}^{\mu,\gamma}\left[\mathbf{1}_K(x_k)+\sum_{j=k+1}^{N}\left(\prod_{i=k}^{j-1}\mathbf{1}_{K'\backslash K}(x_i)\right)\mathbf{1}_K(x_j)\right],$$

$k=0,1,\ldots,N-1$.

From this definition we can infer that $r_{x_0}^{\mu,\gamma}(K,K') = V_0^{\mu,\gamma}(x_0)$. Now consider a recursion operator $T_{f,g}$, parameterized by a one-stage Player I policy $f: X\to\mathcal{A}$ and a one-stage Player II strategy $g: X\times\mathcal{A}\to\mathcal{D}$:

$$T_{f,g}(J)(x) := \mathbf{1}_K(x)+\mathbf{1}_{K'\backslash K}(x)H(x,f(x),g(x,f(x)),J),\quad x\in X, \tag{14}$$

where H is defined in (9). The following result provides a recursive algorithm for computing the functions $V_k^{\mu,\gamma}$.

Lemma 2. *Let* $\mu\in\mathcal{M}_a$, $\gamma\in\Gamma_d$, *and* $V_N^{\mu,\gamma}=\mathbf{1}_K$. *Then for* $k=0,1,\ldots,N-1$, *the following identity holds*

$$V_k^{\mu,\gamma}(x) = T_{\mu_k,\gamma_k}(V_{k+1}^{\mu,\gamma})(x),\quad \forall x\in X. \tag{15}$$

The proof proceeds by minor modifications of previous results in the single-player case (see Lemma 1 of Abate et al., 2008 and Lemma 4 of Summers & Lygeros, 2010), and is omitted.

Next, we will show that under Assumption 1, the operator T defined in (9) preserves suitable measurability properties (thus allowing recursive dynamic programming calculations) and that there exists a one-stage Player I policy and Player II strategy achieving the supremum and infimum in (9). Let $\mathcal{F}$ denote the set of Borel-measurable functions from X to $[0,1]$.

Proposition 3. *If Assumption 1 holds, then*

(a) $\forall J\in\mathcal{F}, T(J)\in\mathcal{F}$;

(b) *For any* $J\in\mathcal{F}$, *there exists a Borel-measurable function* $g^*: X\times\mathcal{A}\to\mathcal{D}$ *such that, for all* $(x,a)\in X\times\mathcal{A}$,

$$g^*(x,a)\in\arg\min_{d\in\mathcal{D}} H(x,a,d,J);$$

(c) *For any* $J\in\mathcal{F}$, *there exists a Borel-measurable function* $f^*: X\to\mathcal{A}$, *such that for all* $x\in X$,

$$f^*(x)\in\arg\max_{a\in\mathcal{A}}\inf_{d\in\mathcal{D}} H(x,a,d,J).$$

Proof. Let $J\in\mathcal{F}$. Define a function $F_J: X\times\mathcal{A}\times\mathcal{D}\to\mathbb{R}$ as $F_J(x,a,d) := H(x,a,d,J)$. From the definition of H, the range of F_J is contained in $[0,1]$. By the Borel-measurability of J and τ, Proposition 7.29 of Bertsekas and Shreve (1978) implies that F_J is Borel-measurable. Furthermore, by Assumption 1 and Fact 3.9 of Nowak (1985), $F_J(x,a,d)$ is continuous in a and d, for each $x\in X$. Now consider a function $\tilde{F}_J(x,a) := \inf_{d\in\mathcal{D}} F_J(x,a,d)$. By the compactness of $\mathcal{D}$ and continuity of F_J in d, this infimum is achieved for each fixed (x,a) (Rudin, 1976). Thus, applying Corollary 1 of Brown and Purves (1973), we have that there exists a Borel-measurable function $g^*: X\times\mathcal{A}\to\mathcal{D}$ for which part (b) holds. Furthermore, by Proposition 7.32 of Bertsekas and Shreve (1978), $\tilde{F}_J$ is continuous in a. Let $F_J^*(x) := \sup_{a\in\mathcal{A}}\tilde{F}_J(x,a) = -\inf_{a\in\mathcal{A}}-\tilde{F}_J(x,a)$. Then, by a repeated application of Corollary 1 of Brown and Purves (1973), there exists a Borel-measurable function $f^*: X\to\mathcal{A}$ such that part (c) holds. By the composition of Borel-measurable functions, this also implies that F_J^* is Borel-measurable.

Finally, it can be observed that $T(J)(x)=\mathbf{1}_K(x)+\mathbf{1}_{K'\backslash K}(x)F_J^*(x)$, $\forall x\in X$. Given that Borel-measurability is preserved under summation and multiplication (see for example Proposition 2.6 of Folland, 1999), $T(J)$ is Borel-measurable. It is also clear that $0\le T(J)\le 1$. Part (a) then follows. □

Now consider the value function $V^* := T^N(\mathbf{1}_K)$. In the following results, it will be shown, through two complementary inequalities, that V^* is in fact equal to the max–min reach–avoid probability $r_{x_0}^*(K,K')$. Furthermore, from the operator T, we also extract player strategies satisfying (10).

Proposition 4. (a) $\forall x_0\in X, T^N(\mathbf{1}_K)(x_0)\le r_{x_0}^*(K,K')$;

(b) *There exists* $\mu^*\in\mathcal{M}_a$ *such that, for any* $\gamma\in\Gamma_d$, $T^N(\mathbf{1}_K)(x_0)\le r_{x_0}^{\mu^*,\gamma}(K,K')$, $\forall x_0\in X$.

Proof. For notational convenience, we define $V_k^* := T^{N-k}(\mathbf{1}_K)$, $k=0,1,\ldots,N$. First, we prove the following claim by backwards induction on k: there exists $\mu_{k\to N}^* = (\mu_k^*,\mu_{k+1}^*,\ldots,\mu_{N-1}^*)\in\mathcal{M}_a$ such that, for any $\gamma_{k\to N}=(\gamma_k,\gamma_{k+1},\ldots,\gamma_{N-1})\in\Gamma_d$, $V_k^*\le V_k^{\mu_{k\to N}^*,\gamma_{k\to N}}$.

Let $\gamma_{k\to N}\in\Gamma_d$ be arbitrary. The case of $k=N$ is trivial. Now assume that this holds for $k=h$. Let $\mu_{h\to N}^*\in\mathcal{M}_a$ be a Player I policy satisfying the induction hypothesis. By Proposition 3(c), there exists a Borel-measurable function $f^*: X\to\mathcal{A}$ such that $f^*(x)\in\arg\max_{a\in\mathcal{A}}\inf_{d\in\mathcal{D}} H(x,a,d,V_h^*)$, $\forall x\in X$. Choose a policy $\mu_{h-1\to N}^* := (f^*,\mu_{h\to N}^*)$. Then by the monotonicity of the operator $T_{f,g}$ and Lemma 2, we have for each $x\in X$:

$$\begin{aligned}
V_{h-1}^{\mu_{h-1\to N}^*,\gamma_{h-1\to N}}(x) &= T_{f^*,\gamma_{h-1}}(V_h^{\mu_{h\to N}^*,\gamma_{h\to N}})(x)\\
&\ge T_{f^*,\gamma_{h-1}}(V_h^*)(x)\\
&= \mathbf{1}_K(x)+\mathbf{1}_{K'\backslash K}(x)H(x,f^*(x),\gamma_{h-1}(x,f^*(x)),V_h^*)\\
&\ge \mathbf{1}_K(x)+\inf_{d\in\mathcal{D}}\mathbf{1}_{K'\backslash K}(x)H(x,f^*(x),d,V_h^*)\\
&= T(V_h^*)(x) = V_{h-1}^*(x).
\end{aligned}$$

The claim then follows by induction. From this, we obtain $\mu_{0\to N}^*\in\mathcal{M}_a$ satisfying $T^N(\mathbf{1}_K)(x_0)=V_0^*(x_0)\le V_0^{\mu_{0\to N}^*,\gamma_{0\to N}}(x_0) = r_{x_0}^{\mu_{0\to N}^*,\gamma_{0\to N}}(K,K')$, $\forall x_0\in X$, $\gamma_{0\to N}\in\Gamma_d$, and hence satisfying statement (b). Furthermore, since $\gamma_{0\to N}$ is arbitrary, $T^N(\mathbf{1}_K)(x_0)\le\inf_{\gamma\in\Gamma_d} r_{x_0}^{\mu_{0\to N}^*,\gamma}(K,K')$, $\forall x_0\in X$. Statement (a) then follows. □

Proposition 5. (a) $\forall x_0\in X, r_{x_0}^*(K,K')\le T^N(\mathbf{1}_K)(x_0)$;

(b) *There exists* $\gamma^*\in\Gamma_d$ *such that, for any* $\mu\in\mathcal{M}_a$, $r_{x_0}^{\mu,\gamma^*}(K,K')\le T^N(\mathbf{1}_K)(x_0)$, $\forall x_0\in X$.

Proof. As before, we define $V_k^* := T^{N-k}(\mathbf{1}_K)$, $k=0,1,\ldots,N$. First, we prove the following claim by backwards induction on k: there exists $\gamma_{k\to N}^* = (\gamma_k^*,\ldots,\gamma_{N-1}^*)\in\Gamma_d$ such that, for any $\mu_{k\to N}=(\mu_k,\ldots,\mu_{N-1})\in\mathcal{M}_a$, $V_k^{\mu_{k\to N},\gamma_{k\to N}^*}\le V_k^*$.

Let $\mu_{k\to N}\in\mathcal{M}_a$ be arbitrary. The case of $k=N$ is trivial. Now assume that this holds for $k=h$. Let $\gamma_{h\to N}^*\in\Gamma_d$ be a Player II strategy satisfying the induction hypothesis. By Proposition 3(b), there exists a Borel-measurable function $g^*: X\times\mathcal{A}\to\mathcal{D}$ such that $g^*(x,a)\in\arg\min_{d\in\mathcal{D}} H(x,a,d,V_h^*)$ for every $x\in X$ and

$a \in \mathcal{A}$. Choose a strategy $\gamma^*_{h-1\to N} := (g^*, \gamma^*_{h\to N})$. Then we have for each $x \in X$:

$$\begin{aligned} V_{h-1}^{\mu_{h-1}\to N,\gamma^*_{h-1\to N}}(x) &= T_{\mu_{h-1},g^*}(V_h^{\mu_{h\to N},\gamma^*_{h\to N}})(x) \\ &\le T_{\mu_{h-1},g^*}(V_h^*)(x) \\ &= \mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x)H(x,\mu_{h-1}(x),g^*(x,\mu_{h-1}(x)),V_h^*) \\ &= \mathbf{1}_K(x) + \inf_{d\in\mathcal{D}} \mathbf{1}_{K'\setminus K}(x)H(x,\mu_{h-1}(x),d,V_h^*) \\ &\le T(V_h^*)(x) = V_{h-1}^*(x). \end{aligned}$$

The claim then follows by induction. From this, we obtain $\gamma^*_{0\to N}$ satisfying $r_{x_0}^{\mu,\gamma^*_{0\to N}}(K,K') = V_0^{\mu,\gamma^*_{0\to N}}(x_0) \le V_0^*(x_0) = T^N(\mathbf{1}_K)(x_0)$, $\forall x_0 \in X$, $\mu \in \mathcal{M}_a$, and hence statement (b). This in turn implies that $r_{x_0}^{\mu}(K,K') = \inf_{\gamma\in\Gamma_d} r_{x_0}^{\mu,\gamma}(K,K') \le T^N(\mathbf{1}_K)(x_0)$, $\forall x_0 \in X$, $\mu \in \mathcal{M}_a$, proving statement (a). □

We are now ready to prove Theorem 1.

Proof. Statement (a) of Theorem 1 follows directly from Propositions 4(a) and 5(a). The Player I policy μ^* and Player II strategy γ^* satisfying statement (b) is provided by Propositions 4(b) and 5(b), respectively. Finally, it can be inferred from the proof of Propositions 4 and 5 that any Player I policy μ^* and Player II strategy γ^* satisfying the conditions in statement (c) is a max–min policy or worst-case strategy, respectively. □

4.2. Specialization to probabilistic safety problem

Consider the probabilistic safety problem defined in (8). Given the connection between the safety and reach–avoid problems through the relation (7), the solution to the probabilistic safety problem can be obtained from a complementary reach–avoid problem. In particular, consider the value function

$$\bar{r}^*_{x_0}(X\setminus S, X) := \inf_{\mu\in\mathcal{M}_a}\sup_{\gamma\in\Gamma_d} r_{x_0}^{\mu,\gamma}(X\setminus S, X), \quad x_0 \in X.$$

From (7) and (8), the max–min safety probability is simply given by

$$p^*_{x_0}(S) = 1 - \bar{r}^*_{x_0}(X\setminus S, X). \tag{16}$$

With minor modifications of Theorem 1, we can show that $\bar{r}^*_{x_0}(X\setminus S, X)$ is computed by the recursion

$$\bar{r}^*_{x_0}(X\setminus S, X) = \bar{T}^N(\mathbf{1}_{X\setminus S})(x_0), \quad x_0 \in X,$$

where the operator $\bar{T}$ is defined as

$$\bar{T}(J)(x) := \mathbf{1}_{X\setminus S}(x) + \inf_{a\in\mathcal{A}}\sup_{d\in\mathcal{D}} \mathbf{1}_S(x)H(x,a,d,J). \tag{17}$$

Combining this with Eq. (16), we obtain the following result.

Theorem 6. *Let $\mathcal{H}$ be a DTSHG satisfying Assumption 1. Let $S \in \mathcal{B}(X)$ be a Borel safe set. Then*

$$p^*_{x_0}(S) = 1 - \bar{T}^N(\mathbf{1}_{X\setminus S})(x_0), \quad \forall x_0 \in X.$$

Similarly as in the reach–avoid problem, max–min safety control policies, as well as worst-case adversary strategies can be derived from the dynamic programming operator given in (17).

4.3. Analytical example

We illustrate the sequence of steps associated with a probabilistic reachability calculation in the context of the jump Markov system example in Section 2.1. In particular, consider a reach–avoid problem in which the objective of Player I is to drive the continuous state into a neighborhood of the origin, while staying within some safe operating region. In this case, the target set and safe set are chosen to be $K := \{q_1, q_2\} \times \left[-\frac{1}{4}, \frac{1}{4}\right]$ and $K' := \{q_1, q_2\} \times [-2, 2]$. In the following, we will solve for the max–min reach–avoid probability and Player I policy over a single stage of the stochastic game ($N = 1$).

Given the DTSHG model, the operator $H(x, a, d, J)$ for a hybrid state $x = (q_1, v)$ can be derived as follows:

$$\begin{aligned} H((q_1,v),a,d,J) &= \int_X J(x')\tau(dx'|(q_1,v),a,d) \\ &= \frac{1}{2}p_1\int_{-1}^{1} J(q_1, 2v+a+d+\eta)d\eta \\ &\quad + \frac{1}{2}(1-p_1)\int_{-1}^{1} J(q_2, 2v+a+d+\eta)d\eta. \end{aligned} \tag{18}$$

For an initial condition $x_0 = (q_1, v_0)$, the max–min reach–avoid probability can be then computed as

$$\begin{aligned} r^*_{(q_1,v_0)}(K,K') &= T(\mathbf{1}_K)(q_1,v_0) \\ &= \begin{cases} 1, & |v_0| \le \frac{1}{4}, \\ 0, & |v_0| > 2, \\ \sup_{a\in\mathcal{A}}\inf_{d\in\mathcal{D}} H((q_1,v_0),a,d,\mathbf{1}_K), & \frac{1}{4} < |v_0| \le 2. \end{cases} \end{aligned} \tag{19}$$

From Eqs. (18) and (19), the analytic expression for the max–min reach–avoid probability in mode q_1 is:

$$r^*_{(q_1,v_0)}(K,K') = \begin{cases} 1, & |v_0| \le \frac{1}{4} \\ \frac{1}{8}, & \frac{1}{4} < |v_0| \le \frac{1}{2} \\ \frac{5}{8} - |v_0|, & \frac{1}{2} < |v_0| \le \frac{5}{8} \\ 0, & |v_0| > \frac{5}{8}. \end{cases}$$

In the course of performing the dynamic programming step in (19), we also obtain a max–min Player I policy μ_0^* in mode q_1 satisfying the sufficient conditions of optimality in (11):

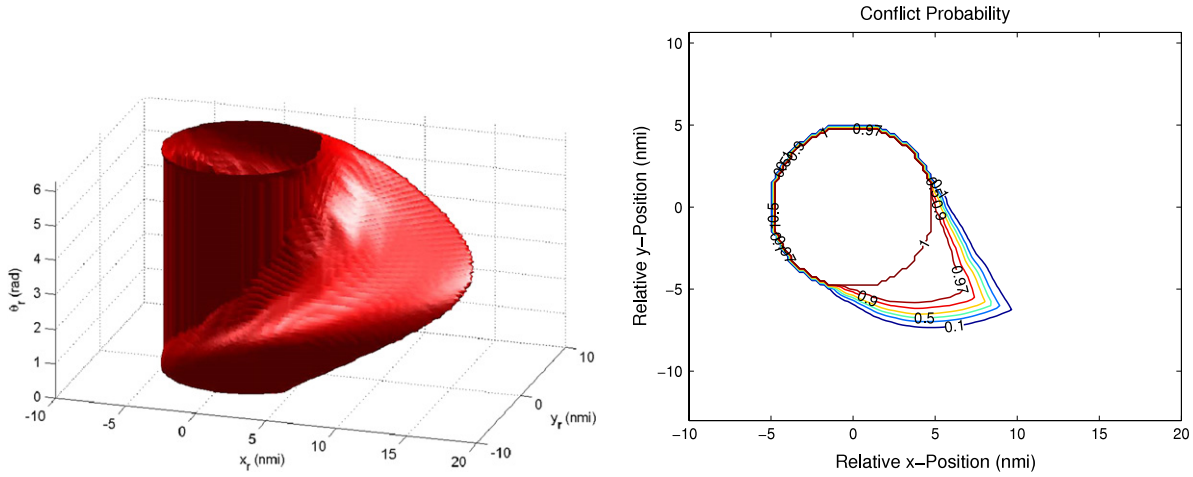
$$\mu_0^*(q_1,v_0) = \begin{cases} -\mathrm{sgn}(v_0), & |v_0| > \frac{1}{2} \\ -2v_0, & |v_0| \le \frac{1}{2}. \end{cases}$$

Using a similar calculation, one can also derive the max–min reach–avoid probability and Player I policy in q_2.

$$r^*_{(q_2,v_0)}(K,K') = \begin{cases} 1, & |v_0| \le \frac{1}{4} \\ \frac{1}{8}, & \frac{1}{4} \le |v_0| \le 2 \\ 0, & |v_0| > 2, \end{cases}$$

$$\mu_0^*(q_2,v_0) = \begin{cases} -\mathrm{sgn}(v_0), & |v_0| > 2 \\ -\frac{1}{2}v_0, & |v_0| \le 2. \end{cases}$$

As one considers more complicated system models, such as in the example of the following section, there may no longer be a closed-form expression for the operator T. This would then require a numerical approximation of the dynamic programming procedure in Theorem 1. In the single-player case, a method is proposed in Abate, Katoen, Lygeros, and Prandini (2010) for a grid-based approximation of the probability map through a discretization of the continuous state space and player input space. However, the



(a) Set of x_0 with conflict probability $\geq 1\%$.

(b) Slice of $\bar{r}^*_{x_0}(X \setminus S, X)$ at $\theta_r = \pi/2$ rad.

Fig. 1. Probability of conflict for pairwise aircraft conflict resolution example.

computational cost of such an approach scales exponentially with the dimensions of the continuous state space and player input spaces, which currently limits application scenarios to problems with relatively low continuous state dimensions (typically $n \leq 4$). Methods for reducing the computational time is a topic of ongoing research Esmaeil Zadeh Soudjani and Abate (2011).

5. Numerical example

In this section, we describe an application of the DTSHG framework to a practical problem in air traffic management, in particular, that of detecting and resolving potential conflicts between pairs of aircraft. For a comprehensive survey of existing methods in this field, the interested reader is referred to Kuchar and Yang (2000). Our approach to this problem involves a combination of worst-case (Tomlin et al., 2002) and probabilistic approaches (Paielli & Erzberger, 1997), namely the intent of one of the aircraft is assumed to be unknown and possibly adversarial, while the wind effects on aircraft trajectory is modeled as stochastic noise. Within this context, conflict resolution then becomes a probabilistic safety problem in which the control task is to maximize the probability of avoiding a collision between two aircraft.

In Paielli and Erzberger (1997), a model for aircraft trajectory perturbation as Gaussian noise was proposed, along with an analytic method for computing the conflict probability. This formed the basis of several probabilistic conflict detection methods which followed Hwang and Seah (2008) and Prandini, Hu, Lygeros, and Sastry (2000). As more detailed trajectory models are considered, with variations to aircraft intent (Yang & Kuchar, 1997) and spatial correlation in wind effects (Hu et al., 2005), closed-form expressions for the conflict probability is often no longer available, requiring the use of numerical computation algorithms. In comparison with previous methods, our approach has the flexibility of being able to treat uncertainty in intent as an adversarial input rather than as a stochastic process, thus offering an interpretation of the conflict probability we compute as the probability of collision under the worst-case behavior of one of the aircraft.

Let $v = (x_r, y_r, \theta_r) \in \mathbb{R}^2 \times [0, 2\pi]$ denote, respectively, the x-position, y-position, and heading of Aircraft 2 in the reference frame of Aircraft 1. By performing a Euler discretization of the kinematics equations in Tomlin et al. (2002) and augmenting the dynamics with a stochastic wind model as in Hu et al. (2005), we obtain the model $v(k+1) = f(v(k), \omega_1(k), \omega_2(k)) + \eta(k)$, where

$$f(v(k), \omega_1(k), \omega_2(k)) := \begin{bmatrix} x_r(k) + \Delta t(-s_1 + s_2\cos(\theta_r(k)) + \omega_1(k) y_r(k)) \\ y_r(k) + \Delta t(s_2 \sin(\theta_r(k)) - \omega_1(k) x_r(k)) \\ \theta_r(k) + \Delta t(\omega_2(k) - \omega_1(k)) \end{bmatrix}. \quad (20)$$

Here Δt is the discretization step, s_i is the speed of aircraft i (assumed to be constant), ω_i is the angular turning rate of aircraft i, taken to be the inputs to the system. The noise vector is given by $\eta = [\eta_1\ \eta_2\ \eta_3]^T$, where $(\eta_1, \eta_2) \sim \mathcal{N}\big(0, 2\sigma_h^2 \Delta t(1 - \exp(-\beta\|(x_r, y_r)\|_2))I\big)$ (details can be found in Hu et al., 2005); and η_3 models process noise acting on the turning rate of either aircraft, with a Gaussian distribution $\eta_3 \sim \mathcal{N}(0, (\sigma_\omega \Delta t)^2)$.

As consistent with common flight maneuvers, we consider a scenario in which each aircraft is allowed to select from among one of three operating modes: straight flight, right turn, or left turn, corresponding to the angular turning rates $\omega_i = 0$, $\omega_i = -\omega$, and $\omega_i = \omega$, respectively. Here, $\omega \in \mathbb{R}$ is assumed to be a constant. The control objective of Aircraft 1 is to avoid a disc D of radius R_c centered on the origin in the (x_r, y_r) plane (corresponding to a loss of minimum separation), subject to the worst-case inputs of Aircraft 2. This can be then viewed as a probabilistic safety problem with the safe set given as $S = D^C \times [0, 2\pi]$. By Theorem 6, the solution to this problem can be obtained from a complementary reach–avoid problem in which the objective of Aircraft 1 is to minimize the worst-case probability of entering the collision set $X \setminus S$, with the value function given by $\bar{r}^*_{x_0}(X \setminus S, X)$.

For our numerical results, we choose a discretization step of $\Delta t = 15$ s, with a time horizon of 2.5 min. The radius of the protected zone is set to $R_c = 5$ nmi. The model parameters are selected as $s_1 = s_2 = 6$ nmi/min, $\omega = 1$ deg/s, $\sigma_h = 0.5$, $\sigma_w = 0.35$, $\beta = 0.1$. The value function is computed using a numerical discretization approach, similar to the one discussed in Abate et al. (2010), on the domain $[-10, 20] \times [-10, 10] \times [0, 2\pi]$, with a grid size of $121 \times 81 \times 73$. We note that for this particular application, the computation of the value function can be performed offline, given wind forecast, and the resulting max–min control policy can be implemented online in lookup table form.

The set of initial conditions x_0 for which the conflict probability is at least 1% (namely, where $\bar{r}^*_{x_0}(X \setminus S, X) \geq 0.01$) is shown in Fig. 1(a). Outside of this set, we have a confidence level of at least 99% of avoiding a collision over a 2.5 min time interval. A slice of the worst-case conflict probability $\bar{r}^*_{x_0}(X \setminus S, X)$ at a relative heading of

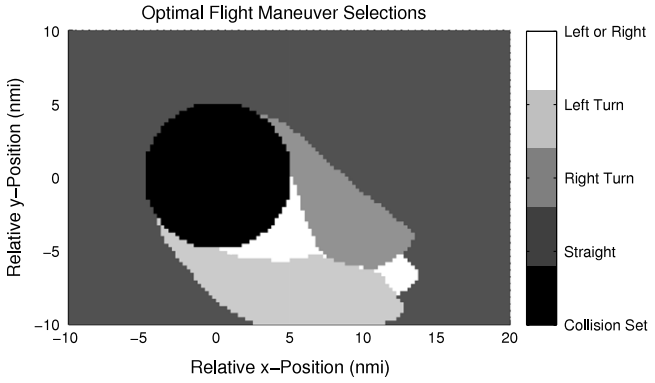


Fig. 2. Max–min control policy at a relative heading of $\theta_r = \pi/2$ rad. The color scale is as follows: black = collision set, dark gray = straight, medium gray = right turn, light gray = left turn, white = either left or right turn.

$\theta_r = \pi/2$ rad is shown in Fig. 1(b). In a conflict detection and resolution algorithm, one can use this probability map to determine the set of states at which to initiate a conflict resolution maneuver (for example where $\bar{r}^*_{x_0}$ exceeds a certain threshold), upon which time the max–min policy μ^* provides a feedback map for selecting flight maneuvers to minimize the conflict probability. A plot of this policy at a relative heading of $\theta_r = \pi/2$ rad is shown in Fig. 2. As can be observed, when the two aircraft are far apart, one can choose to fly straight on the intended course. However, as Aircraft 2 approaches the boundary of the set shown in Fig. 1(a), it becomes necessary for Aircraft 1 to perform an evasive maneuver.

To apply this approach in a large airspace with multiple aircraft, one can obtain the pairwise aircraft conflict probabilities from a probability map such as shown in Fig. 1, for given relative configurations of the aircraft. The air traffic controllers may then define a priority list for trajectory modification, with respect to aircraft pairs whose conflict probabilities exceed a certain threshold.

6. Conclusion

In this article, we discussed a framework for studying probabilistic reachability and safety problems for discrete-time stochastic hybrid systems in a zero-sum stochastic game setting. It was shown that, under certain assumptions on the underlying stochastic kernels and action spaces, there exists a max–min control policy which guarantees a worst-case probability of satisfying the reachability or safety objective, regardless of the adversary strategy. Furthermore, the worst-case probability and the max–min policy can be computed via a dynamic programming recursion.

Some immediate directions for future work are as follows. First, to formally justify approximations of the max–min reach–avoid or safety probabilities through numerical discretization, it would be interesting to establish results on the convergence of the max–min value functions and optimal strategies under appropriate discretization schemes. Based upon existing work in the single-player case (Abate et al., 2010), possible approaches include direct approximation via piecewise-constant functions, or indirect approximation via a finite-state abstraction of the DTSHG model. Second, for application scenarios in which one would like to ensure probabilistic reachability specifications over an extended time horizon, it may be necessary to consider infinite horizon formulations of the safety and reach–avoid problems. Issues here include the convergence of the dynamic programming iterations and the existence of stationary strategies, which may be addressed through adaptation of methods developed for additive cost stochastic games (Kumar & Shiau, 1981; Nowak, 1985). Third, to reduce the conservatism of a max–min approach to reachability problems, one may also consider alternative game formulations with different information patterns. As discussed in Section 2, the existence of equilibrium strategies under a symmetric information pattern is typically assured only under the assumption of randomized policies. This then motivates investigations into methods for efficiently computing and implementing such control policies.

Taking a more long term perspective, the application of the proposed framework to practical problems will require addressing several important challenges. One of the most difficult problems is the development of scalable algorithms for approximating probabilistic reachability computations. A possible approach would be to investigate approximate dynamic programming methods such as the use of adaptive gridding (Esmaeil Zadeh Soudjani & Abate, 2011) or parameterized basis functions (Bertsekas & Tsitsiklis, 1996). Another interesting question is whether the methodology developed here can be used to address multi-objective problems in which one would like to optimize a performance index (e.g. fuel, power consumption), while satisfying a probabilistic reachability specification. Taking a hierarchical view as in Lygeros, Tomlin, and Sastry (1999), one could consider the derivation of design specifications from a reachability computation, for example in terms of the optimality conditions given in Section 4, to serve as constraints for performance optimization. Finally, future research could also investigate probabilistic reachability problems with more complex specifications, such as those expressed in terms of probabilistic computation tree logic (PCTL) (Hansson & Jonsson, 1994). This would require an extension of the proposed methodology to handle temporal objectives (e.g. visiting a sequence of target sets, while remaining safe), possibly through a composition of reach–avoid and safety controllers.

Acknowledgments

This work was supported by the National Science and Engineering Research Council of Canada (NSERC), the "MURI - Frameworks and Tools for High Confidence Design of Adaptive, Distributed Embedded Control Systems" project administered by the Air Force Office of Scientific Research (AFOSR) under Grant FA9550-06-1-0312, the European Commission under the MoVeS project, FP7-ICT-2009-257005, and the European Commission under the Marie Curie grant MANTRAS 249295, and NWO under VENI grant 016.103.020.

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). December.

Abate, A., Katoen, J. P., Lygeros, J., & Prandini, M. (2010). Approximate model checking of stochastic hybrid systems. *European Journal of Control*, *16*(6), 624–641.

Abate, A., Prandini, M., Lygeros, J., & Sastry, S. (2008). Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, *44*(11), 2724–2734.

Ames, A., Sinnet, R., & Wendel, E. (2009). Three-dimensional kneed bipedal walking: a hybrid geometric approach. In R. Majumdar, & P. Tabuada (Eds.), *Lecture notes in computer science*: *vol. 5469. Hybrid systems: computation and control* (pp. 16–30). Springer.

Amin, S., Cardenas, A., & Sastry, S. (2009). Safe and secure networked control systems under denial-of-service attacks. In R. Majumdar, & P. Tabuada (Eds.), *Lecture notes in computer science*: *vol. 5469. Hybrid systems: computation and control* (pp. 31–45). Springer.

Balluchi, A., Benvenuti, L., Di Benedetto, M., Pinello, C., & Sangiovanni-Vincentelli, A. (2000). Automotive engine control and hybrid systems: challenges and opportunities. *Proceedings of the IEEE*, *88*(7), 888–912.

Bect, J., Phulpin, Y., Baili, H., & Fleury, G. (2006). On the Fokker–Planck equation for stochastic hybrid systems: application to a wind turbine model. In *International conference on probabilistic methods applied to power systems* (pp. 1–6).

Bensoussan, A., & Menaldi, J. (2000). Stochastic hybrid control. *Journal of Mathematical Analysis and Applications*, *249*(1), 261–288.

Bertsekas, D. P., & Shreve, S. E. (1978). *Stochastic optimal control: the discrete time case*. New York, NY: Academic Press.

Bertsekas, D. P., & Tsitsiklis, J. (1996). *Neuro-dynamic programming*. Belmont, MA: Athena Scientific.
Breton, M., Alj, A., & Haurie, A. (1988). Sequential Stackelberg equilibria in two-person games. *Journal of Optimization Theory and Applications*, *59*(1), 71–97.
Brown, L. D., & Purves, R. (1973). Measurable selections of extrema. *The Annals of Statistics*, *1*(5), 902–912.
Bujorianu, M. (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.
Esmaeil Zadeh Soudjani, S., & Abate, A. (2011). Adaptive gridding for abstraction and verification of stochastic hybrid systems. In *Proceedings of the 8th international conference on quantitative evaluation of systems*.
Folland, G. B. (1999). *Real analysis: modern techniques and their applications*. New York, NY: John Wiley & Sons.
Ghosh, R., & Tomlin, C. (2004). Symbolic reachable set computation of piecewise affine hybrid automata and its application to biological modelling: Delta–Notch protein signalling. *Systems Biology*, *1*(1), 170–183.
Glover, W., & Lygeros, J. (2004). A stochastic hybrid model for air traffic control simulation. In R. Alur, & G. J. Pappas (Eds.), *Lecture notes in computer science*: *vol. 2993*. *Hybrid systems: computation and control* (pp. 372–386). Springer.
Gonzalez-Trejo, J. I., Hernandez-Lerma, O., & Hoyos-Reyes, L. F. (2002). Minimax control of discrete-time stochastic systems. *SIAM Journal on Control and Optimization*, *41*(5), 1626–1659.
Hansson, H., & Jonsson, B. (1994). A logic for reasoning about time and reliability. *Formal Aspects of Computing*, *6*, 512–535.
Hespanha, J. P. (2004). Stochastic hybrid systems: application to communication networks. In R. Alur, & G. J. Pappas (Eds.), *Lecture notes in computer science*: *vol. 2993*. *Hybrid systems: computation and control* (pp. 47–56). Springer.
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.
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*, *6*(3), 326–340.
Hwang, I., & Seah, C. E. (2008). Intent-based probabilistic conflict detection for the next generation air transportation system. *Proceedings of the IEEE*, *96*(12), 2040–2059.
Kamgarpour, M., Ding, J., Summers, S., Abate, A., Lygeros, J., & Tomlin, C. (2011). Discrete time stochastic hybrid dynamic games: verification & controller synthesis. In *Proceedings of the 50th IEEE conference on decision and control* (pp. 6122–6127). December.
Koutsoukos, X., & Riley, D. (2006). Computational methods for reachability analysis of stochastic hybrid systems. In J. P. Hespanha, & A. Tiwari (Eds.), *Lecture notes in computer science*: *vol. 3927*. *Hybrid systems: computation and control* (pp. 377–391). Springer.
Kuchar, J. K., & Yang, L. C. (2000). A review of conflict detection and resolution modeling methods. *IEEE Transactions on Intelligent Transportation Systems*, *1*(4), 179–189.
Kumar, P. R., & Shiau, T. H. (1981). Existence of value and randomized strategies in zero-sum discrete-time stochastic dynamic games. *SIAM Journal on Control and Optimization*, *19*(5), 617–634.
Kushner, H. J., & Dupuis, P. (1992). *Numerical methods for stochastic control problems in continuous time*. London, UK: Springer-Verlag.
Lincoln, P., & Tiwari, A. (2004). Symbolic systems biology: hybrid modeling and analysis of biological networks. In R. Alur, & G. Pappas (Eds.), *Lecture notes in computer science*: *vol. 2993*. *Hybrid systems: computation and control* (pp. 147–165). Springer.
Lygeros, J., Tomlin, C., & Sastry, S. (1999). Controllers for reachability specifications for hybrid systems. *Automatica*, *35*(3), 349–370.
Maitra, A., & Parthasarathy, T. (1970). On stochastic games. *Journal of Optimization Theory and Applications*, *5*(4), 289–300.
Maitra, A., & Sudderth, W. (1998). Finitely additive stochastic games with Borel measurable payoffs. *International Journal of Game Theory*, *27*(2), 257–267.
Mohajerin Esfahani, P., Chatterjee, D., & Lygeros, J. (2011). On a problem of stochastic reach-avoid set characterization. In *Proceedings of the 50th IEEE conference on decision and control* (pp. 7069–7074). December.
Nowak, A. S. (1985). Universally measurable strategies in zero-sum stochastic games. *The Annals of Probability*, *13*(1), 269–287.
Paielli, R. A., & Erzberger, H. (1997). Conflict probability estimation for free flight. *AIAA Journal of Guidance, Control and Dynamics*, *20*(3), 588–596.
Prajna, S., Jadbabaie, A., & Pappas, G. (2007). A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Transactions on Automatic Control*, *52*(8), 1415–1428.
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.
Rieder, U. (1991). Non-cooperative dynamic games with general utility functions. In T. E. S. Raghavan, T. S. Ferguson, T. Parthasarathy, & O. J. Vrieze (Eds.), *Stochastic games and related topics* (pp. 161–174). Kluwer Academic Publishers.
Rudin, W. (1976). *Principles of mathematical analysis* (3rd ed.). New York, NY: McGraw-Hill.
Sastry, S., Meyer, G., Tomlin, C., Lygeros, J., Godbole, D., & Pappas, G. (1995). Hybrid control in air traffic management systems. In *Proceedings of the 34th IEEE conference on decision and control, vol. 2* (pp. 1478–1483).
Shapley, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences*, *39*(10), 1095–1100.
Summers, S., Kamgarpour, M., Lygeros, J., & Tomlin, C. (2011). A stochastic reach-avoid problem with random obstacles. In *Proceedings of the 14th international conference on hybrid systems: computation and control* (pp. 251–260). ACM.
Summers, S., & Lygeros, J. (2010). Verification of discrete time stochastic hybrid systems: a stochastic reach-avoid decision problem. *Automatica*, *46*(12), 1951–1961.
Tomlin, C., Pappas, G., & Sastry, S. (2002). Conflict resolution for air traffic management: a study in multiagent hybrid systems. *IEEE Transactions on Automatic Control*, *43*(4), 509–521.
Yang, L. C., & Kuchar, L. (1997). Prototype conflict alerting system for free flight. *AIAA Journal of Guidance, Control and Dynamics*, *20*(4), 768–773.

Jerry Ding is a postdoctoral scholar in the Department of Electrical Engineering and Computer Sciences at the University of California, Berkeley. He received his Ph.D. in electrical engineering and computer sciences from UC Berkeley in 2012, and his Bachelor of Science in electrical engineering from the University of Wisconsin-Madison in 2006. He is interested in the development of model-based design and analysis methods for safety-critical systems, with a focus towards control and verification techniques for nonlinear, hybrid, and stochastic systems. His research includes applications of these techniques to problems in autonomous vehicle control, air traffic management, and multi-agent robot motion planning.

Maryam Kamgarpour is a postdoctoral fellow in the department of information technology and electrical engineering at Swiss Federal Institute of Technology, ETH Zurich. Maryam completed her Ph.D. in the department of mechanical engineering at the University of California, Berkeley in 2011. She obtained her bachelor of applied sciences in systems design engineering from University of Waterloo, Canada in 2005. In summer of 2009 she was an intern at NASA Ames Research Center and received the High-Potential Individual Award from NASA Aeronautics.

Maryam's research interests are in modeling, analysis and control of large-scale stochastic hybrid systems. She applies her research to complex engineering systems including air traffic management and the power grid.

Sean Summers received his B.S. degree in Aerospace Engineering in 2004 and his M.S. degree in Mechanical Engineering in 2007, both from the Department of Mechanical and Aerospace Engineering at the University of California, San Diego. He is currently a Ph.D. candidate in the Department of Information Technology and Electrical Engineering at ETH Zurich. His research interests include modeling and control of stochastic hybrid systems, reachability analysis, model predictive control, and applications of control theory in systems biology.

Alessandro Abate is a University Lecturer in the Department of Computer Science at the University of Oxford (UK). He received a Laurea in Electrical Engineering in October 2002 from the University of Padova (IT), an M.S. in May 2004 and a Ph.D. in December 2007, both in Electrical Engineering and Computer Sciences, at UC Berkeley (USA). He has been an International Fellow in the CS Lab at SRI International in Menlo Park (USA), and a Post-Doctoral Researcher at Stanford University (USA), in the Department of Aeronautics and Astronautics. From June 2009 to mid-2013 he has been an Assistant Professor at the Delft Center for Systems and Control, TU Delft–Delft University of Technology (NL). His research interests are in the analysis, verification, and control of probabilistic and hybrid systems, and in their general application over a number of domains, particularly in systems biology and in energy.

John Lygeros is a Professor of Computation and Control at the Swiss Federal Institute of Technology (ETH) Zurich, Switzerland, where he is currently serving as the Head of the Automatic Control Laboratory. He completed a B.Eng. degree in electrical engineering in 1990 and an M.Sc. degree in Systems Control in 1991, both at the Imperial College of Science Technology and Medicine, London, U.K. In 1996 he obtained a Ph.D. degree from the Electrical Engineering and Computer Sciences Department, University of California, Berkeley. During the period 1996–2000 he held a series of research appointments at the National Automated Highway Systems Consortium, Berkeley, the Laboratory for Computer

Science, M.I.T., and the Electrical Engineering and Computer Sciences Department at U.C. Berkeley. Between 2000 and 2003 he was a University Lecturer at the Department of Engineering, University of Cambridge, U.K., and a Fellow of Churchill College. Between 2003 and 2006 he was an Assistant Professor at the Department of Electrical and Computer Engineering, University of Patras, Greece. In July 2006 he joined the Automatic Control Laboratory at ETH Zurich, first as an Associate Professor, and since January 2010 as Full Professor. His research interests include modeling, analysis, and control of hierarchical, hybrid, and stochastic systems, with applications to biochemical networks, automated highway systems, air traffic management, power grids and camera networks. John Lygeros is a Fellow of the IEEE, and a member of the IET and of the Technical Chamber of Greece.

Claire Tomlin is the Charles A. Desoer Professor of EECS at UC Berkeley. From 1998–2007 she was an Assistant, Associate, then Full Professor in Aeronautics and Astronautics at Stanford University, and joined the Berkeley faculty in 2005. Her research is in control and hybrid systems, and she works on applications in robotics, unmanned aerial vehicles, air transportation systems, and systems biology. She received the Eckman Award of the AACC in 2003, a MacArthur Foundation fellowship in 2006, and she became an IEEE Fellow in 2010.