

A Stochastic Games Framework for Verification and Control of Discrete Time Stochastic Hybrid Systems [★]

Jerry Ding ^a, Maryam Kamgarpour, 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*Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands*

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.

Key words: Hybrid Systems, Stochastic Systems, Dynamic Games, Controller Synthesis, Reachability

1 Introduction

In application scenarios ranging from air traffic management [40,44], automotive control [6], systems biology [16,29], to bipedal walking [4], the behavior of the system can be described in terms of a hybrid system abstraction

[★] This work was supported by 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. The corresponding author is Jerry Ding. Tel.: (414) 630-4868. Fax: (510) 642-2718.

Email addresses: jding@eecs.berkeley.edu (Jerry Ding), mkamgar,summers@control.ee.ethz.ch (Maryam Kamgarpour, Sean Summers), a.abate@tudelft.nl (Alessandro Abate), lygeros@control.ee.ethz.ch (John Lygeros), tomlin@eecs.berkeley.edu (Claire Tomlin).

¹ The first two authors contributed equally.

in which the system state evolves both in the discrete and continuous domain. While the discrete state can be used to capture 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) [17,20,21].

For a controlled SHS, the performance of the closed-loop system can be measured in terms of the probability that the system trajectory obeys certain desired specifications. Of interest to safety-critical applications are probabilistic safety and reachability problems in which the control objective is to maximize the probability of remaining within a certain safe set or of reaching a desired target set. In the continuous-time case, a theoretical upper bound on the reachability probability is derived in [13] using Dirichlet forms. The temporal evolution of the

probability density function of the hybrid state has been characterized through generalized Fokker-Planck equations [7]. Optimal control of stochastic hybrid systems is considered in [8] and quasi-variational inequalities based on dynamic programming are derived for the optimal trajectory. An optimal control approach towards reachability analysis is discussed in [25] and [33], in which the solutions of probabilistic safety and reachability problems are derived in terms of the viscosity solutions of appropriate Hamilton-Jacobi-Bellman equations. To address the computational issues associated with probabilistic reachability analysis, the authors in [22] propose a Markov chain approximation of the SHS using methods from [28], while in [36], 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 [3]. These results are generalized in [43] 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 [1] and [42], respectively.

While much of the previous work have studied optimal control formulations of probabilistic safety and reachability 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 is also subject to inputs by an adversary, whose objectives are opposed to that of the control. In particular, generalizing the results in [3] and [43] for a controlled DTSHS, our recent work in [24] 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 [44] and control of networked systems subject to external attacks [5], 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 [24] by providing a thorough exposition of the theoretical results, along with detailed analysis of several examples. Most importantly, we present a detailed proof for the main theorem in [24], 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 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 [3] and [43]. 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 [18] or Stackelberg game formulation [11] 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 [34]. Thus, formal proofs of dynamic programming results requires analysis tools and proof techniques stemming from the field of non-cooperative stochastic games [27,34,38,32,18]. 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 [31,41]. 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 through a simple jump Markov system. In Section 5, we apply our framework to a practical example from 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 [3,43] 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 $\mathcal{Q} := \{q_1, q_2, \dots, q_m\}$, $m \in \mathbb{N}$;
- Dimension of continuous state space $n : \mathcal{Q} \rightarrow \mathbb{N}$: 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 $\tau_v : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times X \times \mathcal{A} \times \mathcal{D} \rightarrow [0, 1]$: a Borel-measurable stochastic kernel on $\mathbb{R}^{n(\cdot)}$ given $X \times \mathcal{A} \times \mathcal{D}$ which assigns to each $x = (q, v) \in X$, $a \in \mathcal{A}$ and $d \in \mathcal{D}$ a probability measure $\tau_v(\cdot | x, a, d)$ on the Borel space $(\mathbb{R}^{n(q)}, \mathcal{B}(\mathbb{R}^{n(q)}))$;
- Discrete state transition kernel $\tau_q : \mathcal{Q} \times X \times \mathcal{A} \times \mathcal{D} \rightarrow [0, 1]$: a Borel-measurable discrete stochastic kernel on \mathcal{Q} given $X \times \mathcal{A} \times \mathcal{D}$ which assigns to each $x \in X$ and $a \in \mathcal{A}$, $d \in \mathcal{D}$ a probability distribution $\tau_q(\cdot | x, a, d)$ over \mathcal{Q} ;
- Reset transition kernel $\tau_r : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times X \times \mathcal{A} \times \mathcal{D} \times \mathcal{Q} \rightarrow [0, 1]$: a Borel-measurable stochastic kernel on $\mathbb{R}^{n(\cdot)}$ given $X \times \mathcal{A} \times \mathcal{D} \times \mathcal{Q}$ which assigns to each $x \in X$, $a \in \mathcal{A}$, $d \in \mathcal{D}$ and $q' \in \mathcal{Q}$ a probability measure $\tau_r(\cdot | x, a, d, q')$ 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 $\mu = (\mu_0, \mu_1, \dots, \mu_{N-1})$ of Borel measurable maps $\mu_k : X \rightarrow \mathcal{A}$, $k = 0, 1, \dots, N-1$. 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 $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_{N-1})$ of Borel measurable maps $\gamma_k : X \times \mathcal{A} \rightarrow \mathcal{D}$, $k = 0, 1, \dots, N-1$. 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 [11]. 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 [31,41]). Furthermore, if one were to consider transition probabilities and 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 [32,38]. However, as shown in [38], 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 hybrid state transition kernel $\tau : \mathcal{B}(X) \times X \times \mathcal{A} \times \mathcal{D} \rightarrow [0, 1]$ which describes the evolution of the hybrid state under the influence of player I and player II inputs:

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

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 $\{x(k), k = 0, \dots, N\}$ 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 $\{x_k, k = 0, \dots, N\}$.

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 $\mathcal{B}(\Omega) := \prod_{k=0}^{N+1} \mathcal{B}(X)$. The evolution of the closed-loop hybrid state trajectory can be described in terms of the transition kernels $\tau^{\mu_k, \gamma_k}(\cdot|x) := \tau(\cdot|x, \mu_k(x), \gamma_k(x, \mu_k(x)))$, $k = 0, \dots, N$. By Proposition 7.28 of [9], 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

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

where $X_0, X_1, \dots, X_N \in \mathcal{B}(X)$ 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 safety and reach-avoid problems, as introduced by [3] and [43] 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, \dots, x_N) reaches K while staying within K' under fixed $\mu \in \mathcal{M}_a$ and $\gamma \in \Gamma_d$ is given by

$$r_{x_0}^{\mu, \gamma}(K, K') := P_{x_0}^{\mu, \gamma}(\{(x_0, \dots, x_N) : \exists j \in \{0, 1, \dots, N\}, (x_j \in K) \wedge (x_i \in K', \forall i \in \{0, 1, \dots, j\})\}) \\ = \sum_{j=0}^N P_{x_0}^{\mu, \gamma}((K' \setminus K)^j \times K \times X^{N-j}), \quad (4)$$

Following a similar procedure as in [43], this probability can be rewritten as

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

where $E_{x_0}^{\mu, \gamma}$ denotes the expectation with respect to the probability measure $P_{x_0}^{\mu, \gamma}$, 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

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

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 $r_{x_0}^*(K, K') := \sup_{\mu \in \mathcal{M}_a} r_{x_0}^{\mu}(K, K'), \forall x_0 \in X$;

- (II) Find a max-min control policy $\mu^* \in \mathcal{M}_a$, whenever it exists, such that $r_{x_0}^*(K, K') = r_{x_0}^{\mu^*}(K, K')$, $\forall x_0 \in X$.
- (III) Find a worst-case adversary strategy $\gamma^* \in \mathcal{M}_a$, whenever it exists, such that $r_{x_0}^*(K, K') = r_{x_0}^{\mu^*, \gamma^*}(K, K')$, $\forall x_0 \in X$.

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 [43], 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

$$\begin{aligned} p_{x_0}^{\mu, \gamma}(S) &:= P_{x_0}^{\mu, \gamma}(\{(x_0, \dots, x_N) : x_k \in S, \forall k\}) \\ &= 1 - r_{x_0}^{\mu, \gamma}(X \setminus S, X). \end{aligned} \quad (7)$$

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

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

Similarly as in the single-player case, both the safety and reach-avoid problems can readily be modified to account for time-varying [1] and stochastic [42] target sets and safe sets. For simplicity of notation, we will focus here on static and deterministic sets.

4 Reach-Avoid Probability Computation

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

The need for continuity assumptions on the stochastic kernel commonly arise in the stochastic game literature (see for example [18,27,31,34]), due to the difficulties in ensuring the measurability of value functions under max-min dynamic programming operations. Following the approach in [34,38], 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 $J : X \rightarrow [0, 1]$ and produces another real-valued function on X :

$$\begin{aligned} T(J)(x) &:= \mathbf{1}_K(x) + \sup_{a \in \mathcal{A}} \inf_{d \in \mathcal{D}} \mathbf{1}_{K' \setminus K}(x) H(x, a, d, J), \\ \text{where } H(x, a, d, J) &:= \int_X J(y) \tau(dy|x, a, d). \end{aligned} \quad (9)$$

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 $T^N = T \circ T \circ \dots \circ T$ (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*

$$r_{x_0}^{\mu^*, \gamma^*}(K, K') \leq r_{x_0}^*(K, K') \leq r_{x_0}^{\mu^*, \gamma^*}(K, K'), \quad (10)$$

$\forall x_0 \in X$, $\mu \in \mathcal{M}_a$, and $\gamma \in \Gamma_d$. In particular, μ^ is a max-min control policy, and γ^* is a worst-case adversary strategy.*

- (c) *Let $J_N = \mathbf{1}_K$, $J_k = T^{N-k}(\mathbf{1}_K)$, $k = 0, 1, \dots, N-1$. If $\mu^* \in \mathcal{M}_a$ is a player I policy which satisfies*

$$\mu_k^*(x) \in \arg \max_{a \in \mathcal{A}} \inf_{d \in \mathcal{D}} H(x, a, d, J_{k+1}), \quad (11)$$

$\forall x \in K' \setminus K$, $k = 0, 1, \dots, N-1$, then μ^ is a max-min control policy. If $\gamma^* = (\gamma_0^*, \gamma_1^*, \dots, \gamma_{N-1}^*) \in \Gamma_d$ is a player II strategy which satisfies*

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

$\forall x \in K' \setminus K$, $a \in \mathcal{A}$, $k = 0, 1, \dots, N-1$, 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

$V_k^{\mu, \gamma} : X \rightarrow [0, 1]$, $k = 0, \dots, N$, defined as

$$\begin{aligned} V_N^{\mu, \gamma}(x_N) &:= \mathbf{1}_K(x_N), \\ V_k^{\mu, \gamma}(x_k) &:= \\ E_{x_k}^{\mu, \gamma} &\left[\mathbf{1}_K(x_k) + \sum_{j=k+1}^N \left(\prod_{i=k}^{j-1} \mathbf{1}_{K' \setminus K}(x_i) \right) \mathbf{1}_K(x_j) \right], \\ k &= 0, 1, \dots, N-1. \end{aligned} \quad (13)$$

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 an one-stage player I policy $f : X \rightarrow \mathcal{A}$ and an one-stage player II strategy $g : X \times \mathcal{A} \rightarrow \mathcal{D}$:

$$\begin{aligned} T_{f, g}(J)(x) &:= \mathbf{1}_K(x) + \\ \mathbf{1}_{K' \setminus K}(x) &H(x, f(x), g(x, f(x)), J), \quad x \in X, \end{aligned} \quad (14)$$

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 $k = 0, 1, \dots, N-1$, the following identity holds*

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

The proof proceeds by minor modifications of previous results in the single-player case (see Lemma 1 of [3] and Lemma 4 of [43]), 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 one-stage player I policy and player II strategy achieving the supremum and infimum in (9). Let \mathcal{F} denote the set of Bore 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 Bore measurable function $g^* : X \times \mathcal{A} \rightarrow \mathcal{D}$ 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 Bore measurable function $f^* : X \rightarrow \mathcal{A}$, 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 $F_J : X \times \mathcal{A} \times \mathcal{D} \rightarrow \mathbb{R}$ 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 [9] implies that F_J is Bore measurable. Furthermore, by Assumption 1 and Fact 3.9 of [34], $F_J(x, a, d)$ 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) [39]. Thus, applying Corollary 1 of [12], we have that there exists a Bore measurable function $g^* : X \times \mathcal{A} \rightarrow \mathcal{D}$ for which part (b) holds. Furthermore, by Proposition 7.32 of [9], \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 [12], there exists a Bore measurable function $f^* : X \rightarrow \mathcal{A}$ such that part (c) holds. By the composition of Bore measurable functions, this also implies that F_J^* is Bore measurable.

Finally, it can be observed that $T(J)(x) = \mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x)F_J^*(x)$, $\forall x \in X$. Given that Bore measurability is preserved under summation and multiplication (see for example Proposition 2.6 of [15]), $T(J)$ is Bore measurable. It is also clear that $0 \leq T(J) \leq 1$. Part (a) then follows. \square

Now consider the dynamic programming value function $J_N := T^N(\mathbf{1}_K)$. In the following two results, it will be shown that J_N is both greater than or equal to, as well as less than or 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) $\forall x_0 \in X, T^N(\mathbf{1}_K)(x_0) \leq r_{x_0}^*(K, K')$;
- (b) *There exists $\mu^* \in \mathcal{M}_a$ such that, for any $\gamma \in \Gamma_d$,*
 $T^N(\mathbf{1}_K)(x_0) \leq r_{x_0}^{\mu^*, \gamma}(K, K'), \forall x_0 \in X$.

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

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

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

The claim then follows by induction. From this, we obtain $\mu_{0 \rightarrow N}^* \in \mathcal{M}_a$ satisfying $T^N(\mathbf{1}_K)(x_0) = J_0(x_0) \leq V_0^{\mu_{0 \rightarrow N}^*, \gamma_{0 \rightarrow N}}(x_0) = r_{x_0}^{\mu_{0 \rightarrow N}^*, \gamma_{0 \rightarrow N}}(K, K')$, $\forall x_0 \in X$, $\gamma_{0 \rightarrow N} \in \Gamma_d$, and hence satisfying statement (b). Furthermore, since $\gamma_{0 \rightarrow N}$ is arbitrary, $T^N(\mathbf{1}_K)(x_0) \leq \inf_{\gamma \in \Gamma_d} r_{x_0}^{\mu_{0 \rightarrow N}^*, \gamma}(K, K')$, $\forall x_0 \in X$. Statement (a) then follows. \square

Proposition 5

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

PROOF. As in the proof of Proposition 4, we define $J_k := T^{N-k}(\mathbf{1}_K)$, $k = 0, 1, \dots, N$. First, we prove the following claim by backwards induction on k : there exists $\gamma_{k \rightarrow N}^* = (\gamma_k^*, \gamma_{k+1}^*, \dots, \gamma_{N-1}^*) \in \Gamma_d$ such that, for any $\mu_{k \rightarrow N} = (\mu_k, \mu_{k+1}, \dots, \mu_{N-1}) \in \mathcal{M}_a$, $V_k^{\mu_{k \rightarrow N}, \gamma_{k \rightarrow N}^*} \leq J_k$.

Let $\mu_{k \rightarrow N} \in \mathcal{M}_a$ be arbitrary. The case of $k = N$ is trivial. Now assume that this holds for $k = h$. Let $\gamma_{h \rightarrow N}^* \in \Gamma_d$ be a player II strategy satisfying the induction hypothesis. By Proposition 3(b), there exists a Borel measurable function $g^* : X \times \mathcal{A} \rightarrow \mathcal{D}$ such that $g^*(x, a) \in \arg \min_{d \in \mathcal{D}} H(x, a, d, J_h)$ for every $x \in X$ and $a \in \mathcal{A}$. Choose a strategy $\gamma_{h-1 \rightarrow N}^* = (g^*, \gamma_{h \rightarrow N}^*)$. Then we have for each $x \in X$:

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

The claim then follows by induction. From this, we obtain $\gamma_{0 \rightarrow N}^* \in \Gamma_d$ satisfying $r_{x_0}^{\mu_{0 \rightarrow N}^*, \gamma_{0 \rightarrow N}^*}(K, K') = V_0^{\mu_{0 \rightarrow N}^*, \gamma_{0 \rightarrow N}^*}(x_0) \leq J_0(x_0) = T^N(\mathbf{1}_K)(x_0)$, $\forall x_0 \in X$, $\mu \in \mathcal{M}_a$, and hence statement (b). This in turn implies that $r_{x_0}^{\mu}(K, K') = \inf_{\gamma \in \Gamma_d} r_{x_0}^{\mu, \gamma}(K, K') \leq T^N(\mathbf{1}_K)(x_0)$, $\forall x_0 \in X$, $\mu \in \mathcal{M}_a$, proving statement (a). \square

We are now ready to prove Theorem 1.

PROOF. Statement (a) of Theorem 1 follows directly from Proposition 4(a) and Proposition 5(a). The player I policy μ^* and player II strategy γ^* satisfying statement (b) is provided by Proposition 4(b) and Proposition 5(b), respectively. Finally, it can be inferred from the proof of Proposition 4 and Proposition 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. \square

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

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

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

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

Combining this with equation (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 Analytic Reach-Avoid 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 $K = \{q_1, q_2\} \times [-\frac{1}{4}, \frac{1}{4}]$ 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:

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

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

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

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

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

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 for optimality in (11):

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

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

$$\begin{aligned} r_{(q_2, v_0)}^*(K, K') &= \begin{cases} 1, & |v_0| \leq \frac{1}{4} \\ \frac{1}{8}, & \frac{1}{4} < |v_0| \leq 2 \\ 0, & |v_0| > 2, \end{cases} \\ \mu_0^*(q_2, v_0) &= \begin{cases} -\text{sgn}(v_0), & |v_0| > 2 \\ -\frac{1}{2}v_0, & |v_0| \leq 2. \end{cases} \end{aligned}$$

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 [2] for a grid-based approximation of the probability map through a discretization of the continuous state space and player input space. However, the 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 [14].

5 Aircraft Conflict Detection and Resolution

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 [26]. Our approach to this problem involves a combination of worst-case [44] and probabilistic approaches [35], namely the intent of one of the aircraft is assumed to be unknown and possibly adversarial, while the wind effects on aircraft trajectory is modelled 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 [35], 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 [23,37]. As more detailed trajectory models are considered, with variations to aircraft intent [45] and spatial correlation in wind effects [22], closed-form expressions for the conflict probability is often no longer available, requiring the use of numerical estimation 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 an Euler discretization of the kinematics equations given in [44] and augmenting the resulting dynamics with a stochastic wind model as described in [22], we obtain the model $v(k+1) = f(v(k), \omega_1(k), \omega_2(k)) + \eta(k)$, where

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

Here Δt is the discretization step, s_i is the speed of aircraft i (assumed to be constant), ω_i is the angular turn-

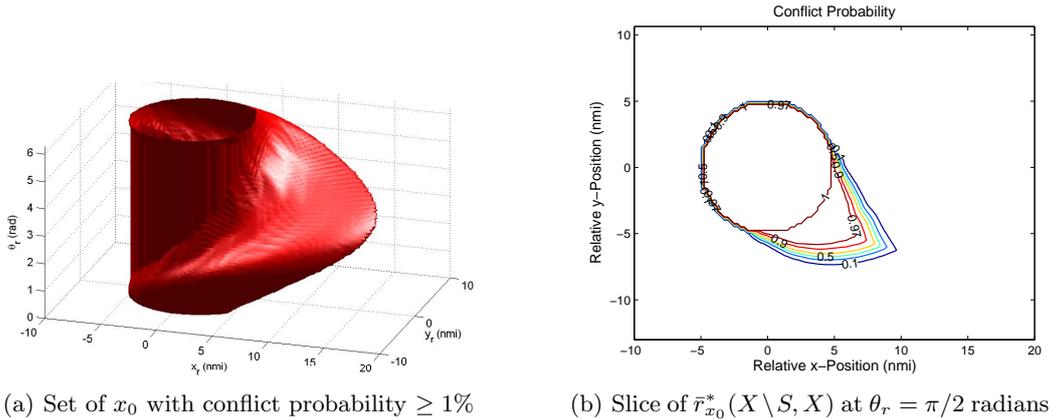


Fig. 1. Probability of conflict for pairwise aircraft conflict resolution example.

ing 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 (η_1, η_2) models spatially correlated wind effects, with a position dependent Gaussian distribution $(\eta_1, \eta_2) \sim \mathcal{N}(0, 2\sigma_h^2 \Delta t (1 - \exp(-\beta \|(x_r, y_r)\|_2)) I)$ (details can be found in [22]); 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 operation 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 sampling time of $\Delta t = 15$ seconds, with a time horizon of 2.5 minutes. 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/sec, $\sigma_h = 0.5$, $\sigma_w = 0.35$, $\beta = 0.1$. The value function is computed using a numerical discretization approach, similar to the one discussed in [2], 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 minute time interval. A slice of the worst-case conflict probability $\bar{r}_{x_0}^*(X \setminus S, X)$ at a relative heading of $\theta_r = \pi/2$ 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 approach the boundary of the set shown in Fig. 1(a), it becomes necessary for aircraft 1 to perform an evasive maneuver.

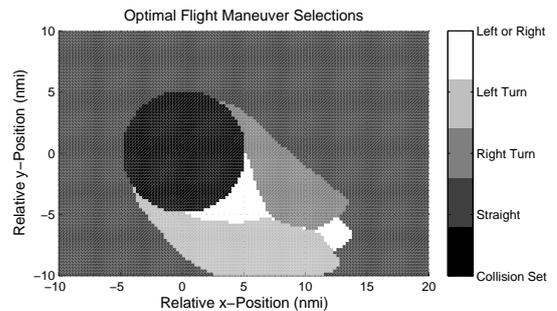


Fig. 2. Max-min control policy at a relative heading of $\theta_r = \pi/2$ rad. The color scale is as follows: Black = collision set, dark gray = straight, medium gray = right turn, light gray = left turn, white = either left or right turn.

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 Figure 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 safety and reachability 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 safety and reachability objectives, 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 safety or reach-avoid 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 [2], 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 [27,34]. 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 is the development of efficient algorithms for approximating probabilistic reachability computations. A possible approach would be to investigate approximate dynamic programming methods such as the use of adaptive gridding [14] or parameterized basis functions [10]. 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 [30], one could consider the derivation of design conditions from a reachability computation, for example in terms of optimality conditions given in Section 4, to serve as constraints for performance optimization. Finally, future research could also investigate into probabilistic reachability problems with more complex specifications, such as expressed in terms of probabilistic computation tree logic (PCTL) [19]. 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.

References

- [1] A. Abate, S. Amin, M. Prandini, J. Lygeros, and S. Sastry. Probabilistic reachability and safe sets computation for discrete time stochastic hybrid systems. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 258–263, December 2006.
- [2] A. Abate, J. P. Katoen, J. Lygeros, and M. Prandini. Approximate model checking of stochastic hybrid systems. *European Journal of Control*, 16(6):624–641, 2010.
- [3] A. Abate, M. Prandini, J. Lygeros, and S. Sastry. Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, 44(11):2724 – 2734, 2008.
- [4] A. Ames, R. Sinnet, and E. Wendel. Three-dimensional kneed bipedal walking: A hybrid geometric approach. In R. Majumdar and P. Tabuada, editors, *Hybrid Systems: Computation and Control*, volume 5469 of *Lecture Notes in Computer Science*, pages 16–30. Springer, 2009.
- [5] S. Amin, A. Cardenas, and S. Sastry. Safe and secure networked control systems under denial-of-service attacks. In R. Majumdar and P. Tabuada, editors, *Hybrid Systems: Computation and Control*, volume 5469 of *Lecture Notes in Computer Science*, pages 31–45. Springer, 2009.
- [6] A. Balluchi, L. Benvenuti, M. Di Benedetto, C. Pinello, and A. Sangiovanni-Vincentelli. Automotive engine control and hybrid systems: Challenges and opportunities. *Proceedings of the IEEE*, 88(7):888–912, 2000.
- [7] J. Bect, Y. Phulpin, H. Baili, and G. Fleury. 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*, pages 1–6, 2006.
- [8] A. Bensoussan and J. Menaldi. Stochastic hybrid control. *Journal of Mathematical Analysis and Applications*, 249(1):261–288, 2000.
- [9] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, New York, NY, 1978.
- [10] D. P. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA., 1996.
- [11] M. Breton, A. Alj, and A. Haurie. Sequential Stackelberg equilibria in two-person games. *Journal of Optimization Theory and Applications*, 59(1):71–97, 1988.
- [12] L. D. Brown and R. Purves. Measurable selections of extrema. *The Annals of Statistics*, 1(5):902–912, 1973.

- [13] M. Bujorianu. Extended stochastic hybrid systems and their reachability problem. In R. Alur and G. Pappas, editors, *Hybrid Systems: Computation and Control*, volume 2993 of *Lecture Notes in Computer Science*, pages 234–249. Springer, 2004.
- [14] S. Esmail Zadeh Soudjani and A. Abate. Adaptive gridding for abstraction and verification of stochastic hybrid systems. In *Proceedings of the 8th International Conference on Quantitative Evaluation of Systems*, September 2011.
- [15] G. B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, New York, NY, 1999.
- [16] R. Ghosh and C. Tomlin. 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, 2004.
- [17] W. Glover and J. Lygeros. A stochastic hybrid model for air traffic control simulation. In R. Alur and G. J. Pappas, editors, *Hybrid Systems: Computation and Control*, volume 2993 of *Lecture Notes in Computer Science*, pages 372–386. Springer, 2004.
- [18] J. I. Gonzalez-Trejo, O. Hernandez-Lerma, and L. F. Hoyos-Reyes. Minimax control of discrete-time stochastic systems. *SIAM Journal on Control and Optimization*, 41(5):1626–1659, 2002.
- [19] H. Hansson and B. Jonsson. A logic for reasoning about time and reliability. *Formal Aspects of Computing*, 6:512–535, 1994.
- [20] J. P. Hespanha. Stochastic hybrid systems: Application to communication networks. In R. Alur and G. J. Pappas, editors, *Hybrid Systems: Computation and Control*, volume 2993 of *Lecture Notes in Computer Science*, pages 47–56. Springer, 2004.
- [21] J. Hu, J. Lygeros, and S. Sastry. Towards a theory of stochastic hybrid systems. In N. Lynch and B. Krogh, editors, *Hybrid Systems: Computation and Control*, volume 1790 of *Lecture Notes in Computer Science*, pages 160–173. Springer, 2000.
- [22] J. Hu, M. Prandini, and S. Sastry. Aircraft conflict prediction in the presence of a spatially correlated wind field. *IEEE Transactions on Intelligent Transportation Systems*, 6(3):326–340, 2005.
- [23] I. Hwang and C. E. Seah. Intent-based probabilistic conflict detection for the next generation air transportation system. *Proceedings of the IEEE*, 96(12):2040–2059, 2008.
- [24] M. Kamgarpour, J. Ding, S. Summers, A. Abate, J. Lygeros, and C. Tomlin. Discrete time stochastic hybrid dynamic games: Verification & controller synthesis. In *Proceedings of the 50th IEEE Conference on Decision and Control*, pages 6122–6127, December 2011.
- [25] X. Koutsoukos and D. Riley. Computational methods for reachability analysis of stochastic hybrid systems. In J. P. Hespanha and A. Tiwari, editors, *Hybrid Systems: Computation and Control*, volume 3927 of *Lecture Notes in Computer Science*, pages 377–391. Springer, 2006.
- [26] J. K. Kuchar and L. C. Yang. A review of conflict detection and resolution modeling methods. *IEEE Transactions on Intelligent Transportation Systems*, 1(4):179–189, 2000.
- [27] P. R. Kumar and T. H. Shiau. Existence of value and randomized strategies in zero-sum discrete-time stochastic dynamic games. *SIAM Journal on Control and Optimization*, 19(5):617–634, 1981.
- [28] H. J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, London, UK, 1992.
- [29] P. Lincoln and A. Tiwari. Symbolic systems biology: Hybrid modeling and analysis of biological networks. In R. Alur and G. Pappas, editors, *Hybrid Systems: Computation and Control*, volume 2993 of *Lecture Notes in Computer Science*, pages 147–165. Springer, 2004.
- [30] J. Lygeros, C. Tomlin, and S. Sastry. Controllers for reachability specifications for hybrid systems. *Automatica*, 35(3):349 – 370, 1999.
- [31] A. Maitra and T. Parthasarathy. On stochastic games. *Journal of Optimization Theory and Applications*, 5(4):289–300, 1970.
- [32] A. Maitra and W. Sudderth. Finitely additive stochastic games with Borel measurable payoffs. *International Journal of Game Theory*, 27(2):257–267, 1998.
- [33] P. Mohajerin Esfahani, D. Chatterjee, and J. Lygeros. On a problem of stochastic reach-avoid set characterization. In *Proceedings of the 50th IEEE Conference on Decision and Control*, pages 7069–7074, December 2011.
- [34] A. S. Nowak. Universally measurable strategies in zero-sum stochastic games. *The Annals of Probability*, 13(1):pp. 269–287, 1985.
- [35] R. A. Paielli and H. Erzberger. Conflict probability estimation for free flight. *AIAA Journal of Guidance, Control and Dynamics*, 20(3):588–596, 1997.
- [36] S. Prajna, A. Jadbabaie, and G. Pappas. A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Transactions on Automatic Control*, 52(8):1415 –1428, 2007.
- [37] M. Prandini, J. Hu, J. Lygeros, and S. Sastry. A probabilistic approach to aircraft conflict detection. *IEEE Transactions on Intelligent Transportation Systems*, 1(4):199–220, 2000.
- [38] U. Rieder. Non-cooperative dynamic games with general utility functions. In T. E. S. Raghavan, T. S. Ferguson, T. Parthasarathy, and O. J. Vrieze, editors, *Stochastic Games and Related Topics*, pages 161–174. Kluwer Academic Publishers, 1991.
- [39] W. Rudin. *Principles of Mathematical Analysis, Third Edition*. McGraw-Hill, New York, NY, 1976.
- [40] S. Sastry, G. Meyer, C. Tomlin, J. Lygeros, D. Godbole, and G. Pappas. Hybrid control in air traffic management systems. In *Proceedings of the 34th IEEE Conference on Decision and Control*, volume 2, pages 1478–1483, December 1995.
- [41] L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953.
- [42] S. Summers, M. Kamgarpour, J. Lygeros, and C. Tomlin. A stochastic reach-avoid problem with random obstacles. In *Proceedings of the 14th International Conference on Hybrid Systems: Computation and Control*, pages 251–260. ACM, April 2011.
- [43] S. Summers and J. Lygeros. Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem. *Automatica*, 46(12):1951 – 1961, 2010.
- [44] C. Tomlin, G. Pappas, and S. Sastry. Conflict resolution for air traffic management: A study in multiagent hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):509–521, 2002.
- [45] L. C. Yang and L. Kuchar. Prototype conflict alerting system for free flight. *AIAA Journal of Guidance, Control and Dynamics*, 20(4):768–773, 1997.