Discrete Time Stochastic Hybrid Dynamical Games: Verification & Controller Synthesis

Maryam Kamgarpour†, Jerry Ding†, Sean Summers, Alessandro Abate, John Lygeros, and Claire Tomlin

***Abstract*— This paper presents a framework for analyzing probabilistic safety and reachability problems for discrete time stochastic hybrid systems in scenarios where system dynamics are affected by rational competing agents. In particular, we consider a zero-sum game formulation of the probabilistic reach-avoid problem, in which the control objective is to maximize the probability of reaching a desired subset of the hybrid state space, while avoiding an unsafe set, subject to the worst-case behavior of a rational adversary. Theoretical results are provided on a dynamic programming algorithm for computing the maximal reach-avoid probability under the worst-case adversary strategy, as well as the existence of a max-min control policy which achieves this probability. The modeling framework and computational algorithm are demonstrated using an example derived from a robust motion planning application.**

I. Introduction

Hybrid dynamical models naturally arise in engineering systems where qualitative behaviors can be abstracted in terms of discrete modes of operation and quantitative behaviors can be characterized in terms of evolution of continuous states. Examples of such systems can be found in a variety of application domains, including air traffic management [1], [2], [3], communication networks [4], systems biology [5], [6], and robotic motion planning [7], [8], [9]. In cases where uncertainties in the system dynamics, for example due to modeling imperfections and environmental disturbances, can be captured using statistical models, the stochastic hybrid system framework [10] provides a powerful tool for analysis and control.

The problem of probabilistic safety for stochastic hybrid systems involves determining the probability that the system trajectory, starting from a given initial condition, will remain inside a safe subset of the discrete and continuous state space (called a hybrid state space) over some given time horizon. On the other hand, the problem of probabilistic reachability involves determining the probability that the system trajectory will reach a desired target set. These problems are of interest, for example, in control and verification problems with safety and target attainability objectives. Here we are interested in a mixture of these two problems, called the *reach-avoid* problem, in which the objective is to characterize the probability that the target set can be attained subject to a safety constraint.

For stochastic hybrid systems, theoretical results on the probabilistic safety and reachability problems are established in [11] and [12]. On the computational side, methods have been proposed for estimating the safety probability through a Markov chain approximation [13] and barrier certificates [14]. A discrete time formulation of these problems is studied in [15], under the framework of Discrete Time Stochastic Hybrid Systems (DTSHS), using techniques from stochastic optimal control [16]. This analysis approach has been generalized to address the reach-avoid problem in [17] for static safe sets and target sets. Extensions to time-varying [18] and stochastic [19] sets have also been considered.

In this paper, we propose a theoretical framework for extending the analysis of the probabilistic reach-avoid problem for DTSHS, as described in [17], to a two-player dynamic game setting. The motivation is that in scenarios where the system dynamics is affected by inputs from rational agents with competing objectives, for example in a network security application [20], or a pursuit-evasion game [21], it is no longer sufficient to simply model adversarial actions as random noise. These scenarios can be more naturally formulated as non-cooperative stochastic games where both the control and adversary are allowed to select rational strategies. Under this framework, we are interested in characterizing the optimal probability of satisfying the reach-avoid objective under the worst-case strategy of a rational adversary, called the *max-min reach-avoid probability*. Also, we would like to find conditions under which there exists an optimal control policy, which we refer to as a *max-min policy*.

The main result of this paper is that under certain standard continuity/compactness assumptions on the stochastic game model [22], [23], the max-min reach-avoid probability is Borel-measurable and can be computed by a suitable dynamic programming algorithm. Furthermore, there exists, under the same set of conditions, a Borel-measurable max-min policy and worst-case adversary strategy.

The contributions of this article are several fold. First, the treatment of the probabilistic reach-avoid problem in a stochastic game setting requires a non-trivial generalization

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 project iFly, FP6-TREN-037180, 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.

M. Kamgarpour is with the Department of Mechanical Engineering, Jerry Ding and Claire Tomlin are with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, USA, {maryamka,jding,tomlin}@eecs.berkeley.edu

Sean Summers and John Lygeros are with the Automatic Control Labratory, ETH Zurich, Switzerland, {summers,lygeros}@control.ee.ethz.ch

Alessandro Abate is with Delft Center for Systems and Control, Delft University of Technology, The Netherlands, a.abate@tudelft.nl

†These authors contributed equally to this work.

of the stochastic optimal control arguments used in [15] and [17]. In particular, measurability properties of the value function are more difficult to establish in a stochastic game setting as opposed to a single-player setting and typically require a stronger set of assumptions on the topological properties of the underlying stochastic kernels and action spaces [24]. Second, given that the pay-off functions for the reach-avoid problem is sum-multiplicative, the result given here can be viewed as an extension of existing results in the stochastic game literature for additive cost games [22], [25], [26]. Third, the existence of a Borel-measurable max-min policy is computationally attractive from a controller synthesis perspective, as Borel-measurable functions can be uniformly approximated by piecewise constant functions (see for example [27]).

The paper is organized as follows. In Section II, we discuss the model for a discrete time stochastic hybrid game. In Section III we formulate the reach-avoid problem in a stochastic game setting. In Section IV, we provide the main result of the paper on the computation of the reach-avoid probability and existence of optimal policies. In Section V we apply the modeling and analysis framework to a motion planning application. Finally, we provide some concluding remarks in Section VI.

II. Discrete Time Stochastic Hybrid Dynamic Game

In this section, we discuss an extension of the DTSHS modeling framework proposed in [15] to allow the stochastic kernels characterizing the hybrid state evolution to depend on the actions of a control and of an adversary. This will be called a Discrete Time Stochastic Hybrid Dynamic Game (DTSHG). Following standard conventions in zero-sum games, we refer to the control as Player I and to the adversary as Player II.

Definition 1 (DTSHG)**.** A discrete-time stochastic hybrid dynamic 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* $\mathcal{Q} := \{q^1, q^2, ..., q^m\}$, $m \in \mathbb{N}$;
- *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 $\mathbb{R}^{n(q)}$. The hybrid state space is given by $X := \bigcup_{q \in \mathcal{Q}} \{q\} \times \mathbb{R}^{n(q)}$;
- *Player I control space* $\mathcal{A}$: a nonempty, compact Borel space;
- *Player II control space* $\mathcal{D}$: a nonempty, compact Borel space;
- *Continuous state transition kernel* $\tau_v(dv'|(q,v),a,d)$: a Borel-measurable stochastic kernel on $\mathbb{R}^{n(q)}$ given $x = (q,v) \in X$, $a \in \mathcal{A}$, and $d \in \mathcal{D}$;
- *Discrete state transition kernel* $\tau_q(q'|(q,v),a,d)$: a discrete stochastic kernel on $\mathcal{Q}$ given $x = (q,v) \in X$, $a \in \mathcal{A}$, and $d \in \mathcal{D}$;
- *Reset transition kernel* $\tau_r(dv'|(q,v),a,d,q')$: a Borel-measurable stochastic kernel on $\mathbb{R}^{n(q')}$ given $x = (q,v) \in X$, $a \in \mathcal{A}$, $d \in \mathcal{D}$, and $q' \in \mathcal{Q}$.

In order to characterize the execution of a DTSHG, it becomes necessary to define how the player I and player II actions are chosen at each time step. It is intuitive that the player whose action is allowed to depend on the choice of action of the other player in general has an advantage in the resulting stochastic game. To be somewhat conservative, we consider an information pattern favorable to Player II. Specifically, at each time step, Player I is allowed to select inputs based upon the current state of the system, while Player II is allowed to select inputs based upon both the system state and the control input of Player I. A mathematical description of this is given below.

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

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

With these definitions, the execution of DTSHG proceeds similarly as in the case of DTSHS, except at the beginning of each time step k, we select player I controls as $a_k = \mu_k(x_k)$ and player II controls as $d_k = \gamma_k(x_k, a_k)$, where $x_k \in X$ is the current state of the DTSHG. Thus, we can define in an analogous fashion as in [15] a stochastic kernel $\tau(dx'|x,a,d)$ which describes the evolution of the hybrid state under player I and player II controls. Let the hybrid state be denoted as $x = (q,v) \in X$, then

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

For a given initial condition $x_0 \in X$, player I policy $\mu \in \mathcal{M}_a$, and player II strategy $\gamma \in \Gamma_d$, the closed-loop execution of the DTSHG is then specified through the following definition.

Definition 4 (DTSHG Execution)**.** Let $\mathcal{H}$ be a DTSHG and $N \in \mathbb{N}$ be a finite time horizon. A stochastic process $\{x_k,\ k = 0, ..., N\}$ with values in X is an execution of $\mathcal{H}$ associated with an initial condition $x_0 \in X$, a player I policy $\mu \in \mathcal{M}_a$, and a player II strategy $\gamma \in \Gamma_d$, if its sample paths are obtained according to Algorithm II.1.

It can be observed that the execution resulting from Algorithm II.1 is a time inhomogeneous stochastic process on the sample space $\Omega = X^{N+1}$, endowed with the canonical product topology $\mathcal{B}(\Omega) := \prod_{k=1}^{N+1} \mathcal{B}(X)$, where $\mathcal{B}(\cdot)$ denotes the Borel σ-algebra on a topological space. For notational conveniences, we define $\tau^{\mu_k,\gamma_k}(\cdot|x) := \tau(\cdot|x, \mu_k(x), \gamma_k(x, \mu_k(x)))$ as the closed-loop hybrid state transition kernel at time k, under given choices of $\mu \in \mathcal{M}_a$ and $\gamma \in \Gamma_d$. For a fixed initial condition $x_0 \in X$, the stochastic kernels τ^{μ_k,γ_k}, $k = 0, 1, .., N$ induce an unique

Algorithm II.1 DTSHG Execution

Input Initial condition $x_0 \in X$, player I policy $\mu \in \mathcal{M}_a$, player II strategy $\gamma \in \Gamma_d$
Output Sample Path $\{x_k, \quad k = 0, ..., N\}$
Set $k = 0$;
while $k < N$ **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} according to $\tau(\cdot|x_k, a_k, d_k)$;
Increment k;
end while

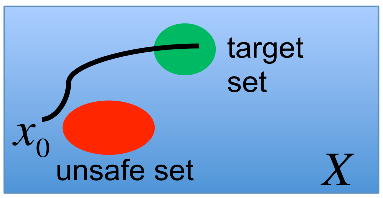

Fig. 1. The probabilistic reach-avoid problem is concerned with finding the probability that the state trajectory, starting from a given initial condition $x_0 \in X$, will reach a target set while avoiding an unsafe set.

probability measure $P^{\mu,\gamma}_{x_0}$ on Ω (see Proposition 7.28 of [16]). In the next section, we describe how this probability measure allows us to quantify the probability of satisfying the reach-avoid objective for player I.

III. Probabilistic Reach-Avoid Problem for DTSHG

Using the modeling framework of a DTSHG, we consider a stochastic game formulation of the probabilistic reach-avoid problem in which player I (the control) has the objective of steering the hybrid state into a desired target set, while avoiding an unsafe set (as illustrated in Fig. 1), and player II (the adversary) has the opposing objective of steering the state into the unsafe set or preventing the state from reaching the target set. This scenario can arise, for example, in a robust control application where we would like design a feedback controller to steer the system state into a neighborhood of an operating point, subject to state constraints on the closed-loop trajectory and disturbances acting on the system dynamics. In contrast with the single-player case, as addressed in [17], an optimal control policy for the DTSHG needs to account for the worst-case strategy of the adversary.

In the following, we proceed to give a more precise statement of the problem. As in [17], we assume that the Borel sets $K, K' \in \mathcal{B}(X)$ are given as the target set and safe set, respectively, with $K \subseteq K'$. For a given initial condition $x_0 \in X$, player I policy $\mu \in \mathcal{M}_a$, and player II strategy $\gamma \in \Gamma_d$, the probability that the execution $(x_0, x_1, ..., x_N)$ of a DTSHG reaches K at some time $j = 0, 1, \ldots, N$ in the horizon of interest, while staying inside K' at all previous times $i = 0, 1, \ldots, j$ is given by

$$
\begin{aligned}
r^{\mu,\gamma}_{x_0}(K, K') &:= P^{\mu,\gamma}_{x_0}\left(\bigcup_{j=0}^{N} (K' \setminus K)^j \times K \times X^{N-j}\right) \\
&= \sum_{j=0}^{N} P^{\mu,\gamma}_{x_0}((K' \setminus K)^j \times K \times X^{N-j}), \quad (2)
\end{aligned}
$$

where the second equality follows by the fact that the union is disjoint. By Proposition 7.28 of [16], this probability can be computed as

$$
\begin{aligned}
&r^{\mu,\gamma}_{x_0}(K, K') = \quad (3) \\
&E^{\mu,\gamma}_{x_0}\left[\mathbf{1}_K(x_0) + \sum_{j=1}^{N}\left(\prod_{i=0}^{j-1} \mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_j)\right],
\end{aligned}
$$

which is analogous to the sum-multiplicative cost given in [17] for a DTSHS. Now define the worst-case reach-avoid probability under a choice of Markov policy μ as

$$r^{\mu}_{x_0}(K, K') = \inf_{\gamma \in \Gamma_d} r^{\mu,\gamma}_{x_0}(K, K'). \quad (4)$$

Our control objective is then to maximize this worst-case probability over the set of Markov policies:

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

$$r^{*}_{x_0}(K, K') := \sup_{\mu \in \mathcal{M}_a} r^{\mu}_{x_0}(K, K'), \; x_0 \in X; \quad (5)$$

(II) Find a max-min policy $\mu^* \in \mathcal{M}_a$, whenever it exists, such that $r^{*}_{x_0}(K, K') = r^{\mu^*}_{x_0}(K, K')$, $\forall x_0 \in X$.

IV. Solution Approach and Implications

In this section, we state a result regarding the computation of the max-min reach-avoid probability through dynamic programming and the existence of max-min policies. To ensure that the desired measurability properties are preserved in a dynamic programming recursion, we will require the following additional assumptions on the stochastic kernels of the DTSHG, as inspired by [22], [23].

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

It should be noted that we only assume continuity of the stochastic kernels in the actions of Player I and Player II, but not necessarily in the system state. Thus, our Borel-measurable model still allows for stochastic hybrid systems where 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 above assumptions are clearly satisfied under the discrete topology on $\mathcal{A}$ and $\mathcal{D}$. Also, if $\tau_v(\cdot|(q,v),a,d)$ has a density function $f_v(v'|(q,v),a,d), v' \in \mathbb{R}^{n(q)}$ for every $q \in \mathcal{Q}$, and $f_v(v'|(q,v),a,d)$ is continuous in a and d, it can be checked that the assumption for τ_v is satisfied. A similar condition can also be formulated for the reset kernel τ_r.

Now define a dynamic programming operator T which maps a Borel-measurable function $J : X \to [0,1]$ to a function $T(J) : X \to [0,1]$ as defined by

$$T(J)(x) = \sup_{a\in\mathcal{A}} \inf_{d\in\mathcal{D}} \mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x) H(x,a,d,J), \quad (6)$$

where $H(x,a,d,J) := \int_X J(x')\tau(dx'|x,a,d)$.

The main result of the paper is as follows.

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 (6). 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 player II strategy $\gamma^* \in \Gamma_d$ satisfying

$$r^{\mu,\gamma^*}_{x_0}(K,K') \le r^*_{x_0}(K,K') \le r^{\mu^*,\gamma}_{x_0}(K,K'), \quad (7)$$

for every $x_0 \in X$, $\mu \in \mathcal{M}_a$, and $\gamma \in \Gamma_d$. In particular, μ^* is a max-min policy for player I.

Aside from providing us with a dynamic programming algorithm for computing the max-min reach-avoid probability, this result also gives a more precise characterization of the max-min policy. In particular, we have by (7) that if the control were to select the max-min policy μ^* and the adversary were to deviate from the worst-case strategy γ^*, then the reach-avoid probability will be at least $r^*_{x_0}(K,K')$. On the other hand, if the control were to deviate from the max-min policy μ^* and the adversary were to choose the worst-case strategy γ^*, then the reach-avoid probability will be at most $r^*_{x_0}(K,K')$. Thus, μ^* can be interpreted as a robust control policy which optimizes a worst-case performance index.

Due to space limitations, the proof of the theorem is omitted. Instead we will highlight here the main points of the proof. The interested reader is referred to REF:TechReport for further details. First, we can show in a similar manner as in [15] and [17] that the reach-avoid probability $r^{\mu,\gamma}_{x_0}(K,K')$ under fixed $\mu \in \mathcal{M}_a$ and $\gamma \in \Gamma_d$ is computed by a recursive formula. Second, we can prove a max-min selection theorem for T, as an application of [24] and [28], showing that the operator T preserves measurability properties and that there exists Borel-measurable selectors which achieve the supremum and infimum in (6). Finally, using the recursive formula for $r^{\mu,\gamma}_{x_0}(K,K')$ and the selection theorem for T, we can show that $T^N(\mathbf{1}_K)$ simultaneously upper bounds and lower bounds $r^*_{x_0}(K,K')$ and that there exist a player I policy and player II strategy which satisfy (7). The last step can be seen as an extension of the dynamic programming results for additive cost stochastic games [22], [25], [26] to the sum-multiplicative case.

On a computational note, the dynamic programming recursion in Theorem 1 can be carried out in an approximate fashion through a discretization of the continuous state space and player control spaces. Specifically, suppose that an analytic characterization of the hybrid state transition kernel τ is available (for example as a probability density function over the continuous state space within each mode), then for each grid point $x_g \in X$, and discretized inputs $a_g \in \mathcal{A}$ and $d_g \in \mathcal{D}$, the operator $H(x_g,a_g,d_g,J)$ in (6) can be computed by integration of J over X, under the probability measure $\tau(\cdot|x_g,a_g,d_g)$. This then provides a piece-wise constant approximation of the value function through a discrete grid representation. In [29], this type of discretization schume is shown to converge uniformly to the maximal safety probability for a DTSHS, at a rate that is linear in the grid size parameter. For the case where an analytic expression for τ is not available, Monte Carlo simulation may be used to approximate the transition probabilities, as discussed in [17]. We note however that the computational complexity of this approach does scale exponentially with the dimension of the continuous state space. Finding methods to reduce this computational complexity is a topic of ongoing research [30].

V. Computational Example

Here we provide a practical example from the domain of robust motion planning to illustrate the modeling framework and solution approach discussed previously. Specifically, we consider a target tracking application where the control objective is to drive an autonomous quadrotor helicopter to a hover region over a moving ground vehicle within finite time, while satisfying certain velocity constraints. This problem has been addressed in [31] using a continuous time robust control framework, and experimental tests have been performed on the Stanford Testbed of Autonomous Rotorcraft for Multi-Agent Control (STARMAC) [32].

Through experimental trials, the position-velocity dynamics of the quadrotor is found to be well-approximated by a double integrator model in the planar axis x and y, with some added disturbance terms to account for the movement of the ground vehicle and the effects of model uncertainties and actuator noise. Using the DTSHG framework, we will assume a probabilistic model for the noise entering into the quadrotor dynamics, while allowing the ground vehicle to choose inputs rationally within its acceleration limits. More specifically, consider the following stochastic model for the relative motion between the quadrotor and the ground vehicle:

$$\begin{aligned} x_1(k+1) =& x_1(k) + \Delta t x_2(k) + \\ & \frac{\Delta t^2}{2}(g\sin(\phi(k)) + d_x(k)) + \eta_1(k) \\ x_2(k+1) =& x_2(k) + \Delta t(g\sin(\phi_k) + d_x(k)) + \eta_2(k) \end{aligned}$$

$$y_1(k+1) = y_1(k) + \Delta t y_2(k) + \frac{\Delta t^2}{2}(g\sin(-\theta(k)) + d_y(k)) + \eta_3(k)$$
$$y_2(k+1) = y_2(k) + \Delta t(g\sin(-\theta(k)) + d_y(k)) + \eta_4(k),$$

where x_1, x_2, y_1, y_2 are the position and velocity of the quadrotor relative to the ground vehicle in the $x-$axis and $y-$axis respectively, Δt is the discretization step, ϕ is the quadrotor roll angle command, θ is the quadrotor pitch angle command, and g is the gravitational constant.

The disturbance parameters in this model include d_x and d_y, which are the accelerations of the ground vehicle in the x and y directions, as well as η_i, $i = 1, \ldots, 4$, which represent the model uncertainties and actuator noise. Given that the ground vehicle may be a rational agent, we take d_x and d_y to be the inputs of player II. On the other hand, we model η_i as normally distributed according to $\eta_i \sim \mathcal{N}(0, ({\sigma_i}^2)\Delta t^2)$.

In order to complete the description of the DTSHG model, we note that the inputs ϕ and θ are selected from a quantized input range due to digital implementation. These quantization levels can be viewed as the discrete states of the system, resulting in a discrete time switched system.

For the target tracking application, the target set is chosen to be a square-shaped hover region centered on the ground vehicle, with some tolerance on the relative velocity. In (x_1, x_2) coordinates, this is specified as

$$K_x = [-0.2, 0.2]m \times [-0.2, 0.2]m/s.$$

The safety constraint in this case is chosen to be a bound on the permissible relative position and velocity. In (x_1, x_2) coordinates, this is specified as

$$K'_x = [-1.2, 1.2]m \times [-1, 1]m/s.$$

The corresponding target set K_y and safe set K'_y in the y direction are specified identically. The target and safe sets in two dimensions are then defined as $K = K_x \times K_y$ and $K' = K'_x \times K'_y$ respectively. Under a stochastic game formulation of the problem, the objective of the quadrotor (player I) is to reach the hover region K while satisfying the state and velocity constraint K', subject to the worst-case acceleration inputs (d_x, d_y) of the ground vehicle.

Given the problem description, we decouple the reach-avoid probability computation into two independent calculations in the (x_1, x_2) and (y_1, y_2) coordinates. For the numerical results to be shown here, the roll and pitch commands are chosen from the range $[-10°, 10°]$, quantized at $2.5°$ intervals, while the acceleration bounds for d_x and d_y are chosen to be $[-.4, .4]$ m/s^2 and are discretized at $0.1\frac{m}{s}$ intervals for numerical computation. The variance of the noise parameters is set to be $\sigma_i = 0.4$. The time step is set to be $\Delta t = 0.1$s. The time horizon is chosen to be one second ($N = 10$).

Using the dynamic programming algorithm discussed in Section IV, we compute the max-min reach-avoid probability for the quadrotor over the safety constraint set K'_x in (x_1, x_2) coordinates, using a discrete grid of 61×41 nodes. The result is shown in Fig. 2. The corresponding contours of this probability map are plotted in Fig. 3, with the target set K shown in the center with probability contour one. Due to the symmetry of the problem, only the results for the $x-$axis are shown.

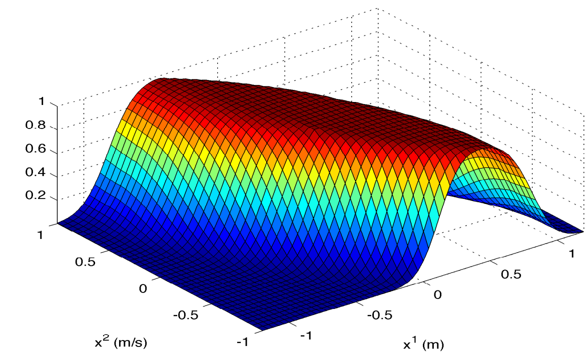

Fig. 2. Probability of reach-avoid for the relative position and velocity of the quadrotor with respect to the ground vehicle.

An interpretation of these results can be given as follows. Suppose we initialize the quadrotor at a relative x-position $x_1(0)$ and relative x-velocity $x_2(0)$ within the 0.8 probability contour in Fig. 2, namely where $r^*_{(x_1(0), x_2(0))}(K, K') \geq 0.8$. Then by Theorem 1, if the quadrotor were to select its roll angle commands according to the max-min control policy μ^* over a time interval of one second, then it will safely reach the hover region with a probability of at least 80%, regardless of the choice of acceleration inputs by the ground vehicle. Thus, the set of states $\left\{(x_1, x_2) : r^*_{(x_1, x_2)}(K, K') \geq 0.8\right\}$ form the set of feasible initial conditions for which there exists a feedback policy satisfying the target tracking specifications with at least 80% probability over the time horizon of interest.

In comparison with a deterministic reachability approach, such as considered in [31], the probabilistic reachability analysis discussed here provides a measure of confidence in the system performance in cases where hard bounds on the noise parameters is not available, but a statistical model can be obtained for the noise parameters through experimental data (for example as a Gaussian distribution). On the other hand, when bounds for some of the disturbances are available (in this case the acceleration inputs of the ground vehicle), the DTSHG model still allows us to determine the performance of the system under the worst-case behavior of these disturbances, as well as to construct a control policy which optimizes this worst-case performance.

VI. Conclusion

In this work, we described a framework for extending the study of probabilistic safety and reachability problems for discrete time stochastic hybrid systems to a stochastic game setting in which the evolution of the system state is affected by the decisions of two rational agents. The probabilistic reach-avoid problem is formulated within this framework as a zero-sum game between a control and an adversary. A solution to this problem is provided in the form of a dynamic

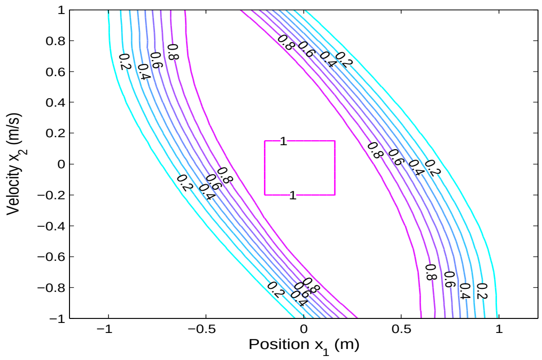

Fig. 3. Reach-avoid probability contours for the target tracking game.

programming algorithm for computing the max-min reach-avoid probability and an existence and synthesis result for the max-min policy. Some directions for future work include tractable approximation schemes for the reach-avoid probability, extensions to infinite horizon reachability problems, and investigation of alternative information patterns between the control and the adversary.

References

[1] S. Sastry, G. Meyer, C. Tomlin, J. Lygeros, D. Godbole, and G. Pappas, "Hybrid control in air traffic management systems," in *IEEE Conference on Decision and Control*, vol. 2, 1995, pp. 1478–1483.

[2] C. Tomlin, G. Pappas, and S. Sastry, "Conflict resolution for air traffic management: A study in multiagent hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 509–521, 2002.

[3] M. Soler, A. Olivares, and E. Staffed, "Hybrid Optimal Control Approach to Commercial Aircraft Trajectory Planning," *Journal of Guidance, Control and Dynamics*, vol. 33, no. 3, pp. 985–991, 2010.

[4] J. P. Hespanha, "Stochastic hybrid systems: Application to communication networks," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, R. Alur and G. J. Pappas, Eds. Springer Berlin / Heidelberg, 2004, vol. 2993, pp. 47–56.

[5] 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*, vol. 1, no. 1, pp. 170–183, 2004.

[6] P. Lincoln and A. Tiwari, "Symbolic systems biology: Hybrid modeling and analysis of biological networks," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, R. Alur and G. Pappas, Eds. Springer, 2004, pp. 147–165.

[7] A. Ames, R. Sinnet, and E. Wendel, "Three-dimensional kneed bipedal walking: A hybrid geometric approach," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, R. Majumdar and P. Tabuada, Eds. Springer, 2009, pp. 16–30.

[8] J. Gillula, G. Hoffmann, H. Huang, M. Vitus, and C. Tomlin, "Applications of hybrid reachability analysis to robotic aerial vehicles," *International Journal of Robotics Research*, vol. 30, no. 3, pp. 335–354, 2011.

[9] E. Frazzoli, M. Dahleh, and E. Feron, "Robust hybrid control for autonomous vehicle motion planning," in *IEEE Conference on Decision and Control*, vol. 1, 2000, pp. 821–826.

[10] J. Hu, J. Lygeros, and S. Sastry, "Towards a theory of stochastic hybrid systems," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, N. A. Lynch and B. H. Krogh, Eds. Springer, 2000, vol. 1790, pp. 160–173.

[11] M. Bujorianu, "Extended stochastic hybrid systems and their reachability problem," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, R. Alur and G. Pappas, Eds. Springer Berlin / Heidelberg, 2004, vol. 2993, pp. 234–249.

[12] X. D. Koutsoukos and D. Riley, "Computational methods for reachability analysis of stochastic hybrid systems," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, J. P. Hespanha and A. Tiwari, Eds. Springer, 2006, pp. 377–391.

[13] M. Prandini and J. Hu, "Application of reachability analysis for stochastic hybrid systems to aircraft conflict prediction," *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 913 –917, april 2009.

[14] S. Prajna, A. Jadbabaie, and G. Pappas, "A framework for worst-case and stochastic safety verification using barrier certificates," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1415 –1428, aug. 2007.

[15] A. Abate, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems," *Automatica*, vol. 44, no. 11, pp. 2724–2734, November 2008.

[16] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific, February 2007.

[17] S. Summers and J. Lygeros, "Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem," *Automatica*, vol. 46, no. 12, pp. 1951 – 1961, 2010.

[18] A. Abate, S. Amin, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability and safe sets computation for discrete time stochastic hybrid systems," in *IEEE Conference on Decision and Control*, pp. 258–263.

[19] S. Summers, M. Kamgarpour, J. Lygeros, and C. Tomlin, "A Stochastic Reach-Avoid Problem with Random Obstacles," in *Hybrid Systems: Computation and Control*. ACM, 2011, pp. 251–260.

[20] S. Amin, A. Cardenas, and S. Sastry, "Safe and secure networked control systems under denial-of-service attacks," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, R. Majumdar and P. Tabuada, Eds. Springer Berlin / Heidelberg, 2009, vol. 5469, pp. 31–45.

[21] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory*. Society for Industrial Mathematics, 1999.

[22] A. S. Nowak, "Universally measurable strategies in zero-sum stochastic games," *The Annals of Probability*, vol. 13, no. 1, pp. pp. 269–287, 1985.

[23] U. Rieder, "Non-cooperative dynamic games with general utility functions," in *Stochastic Games and Related Topics*, T. Raghavan, T. S. Ferguson, T. Parthasarathy, and O. J. Vrieze, Eds. Kluwer Academic Publishers, 1991, pp. 161 –174.

[24] A. Nowak, "On measurable minimax selectors," *Journal of Mathematical Analysis and Applications*, vol. 366, no. 1, pp. 385–388, 2010.

[25] 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*, vol. 19, no. 5, pp. pp. 617–634, 1981.

[26] 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*, vol. 41, no. 5, pp. 1626–1659, 2002.

[27] G. B. Folland, *Real Analysis*. John Wiley & Sons, 1999.

[28] L. D. Brown and R. Purves, "Measurable selections of extrema," *The Annals of Statistics*, vol. 1, no. 5, pp. pp. 902–912, 1973.

[29] A. Abate, J. Katoen, J. Lygeros, and M. Prandini, "Approximate model checking of stochastic hybrid systems," *European Journal of Control*, no. 6, pp. 624–641, 2010.

[30] S. Esmaeil Sadegh 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*, Aachen, DE, September 2011.

[31] J. Ding, E. Li, H. Huang, and C. J. Tomlin, "Reachability-based synthesis of feedback policies for motion planning under bounded disturbances," in *2011 IEEE International Conference on Robotics and Automation (ICRA)*, may 2011, pp. 2160 –2165.

[32] G. Hoffmann, H. Huang, S. Waslander, and C. J. Tomlin, "Quadrotor helicopter flight dynamics and control: Theory and experiment," in *AIAA Conference on Guidance, Navigation and Control*, Aug. 2007.