Optimal Control of Partially Observable Discrete Time Stochastic Hybrid Systems for Safety Specifications*

Jerry Ding[1], Alessandro Abate[2], and Claire Tomlin[1]

Abstract— **This paper describes a theoretical framework for the design of controllers to satisfy probabilistic safety specifications for partially observable discrete time stochastic hybrid systems. We formulate the problem as a partial information stochastic optimal control problem, in which the objective is to maximize the probability that the state trajectory remains within a given safe set in the hybrid state space, using observations of the history of inputs and outputs. It is shown that this optimal control problem, which has a multiplicative payoff structure, is equivalent to a terminal payoff problem when the state space is augmented with a binary random variable capturing the safety of past state evolution. This allows us to derive a sufficient statistic for the probabilistic safety problem as a set of Bayesian filtering equations updating a conditional distribution on the augmented state space, as well as an abstract dynamic programming algorithm for computing the maximal probability of safety and an optimal control policy.**

I. INTRODUCTION

For safety-critical applications such as air traffic management [1], automated highway systems [2], and autonomous vehicle control [3], the designs of feedback controllers are often required to satisfy stringent safety specifications on the closed-loop system behavior, as determined by a combination of industry standards and government regulations. The problem of meeting these specifications is complicated by the numerous sources of uncertainties arising within a practical setting, including both environmental disturbances and measurement noise. As a modeling framework, stochastic hybrid systems [4], [5] provide a mathematical formalism for reconciling discrete and continuous abstractions of system behavior, while allowing for a probabilistic description of uncertainty. Within this context, a safety control problem can be formulated to design an observation-based feedback policy maximizing the probability that the state trajectory remains within a safe subset of the hybrid state space.

To illustrate this in terms of a concrete example, one can consider the heating, ventilation, and air conditioning (HVAC) system found in large scale commercial buildings. This system commonly features a complex network of air handling units, boilers, and chillers, with both discrete and continuous elements. In particular, while the room temperatures themselves are inherently continuous and are governed by thermodynamic laws, the actuators often include switching devices such as valves, dampers, and pumps. Hybrid system models have been previously proposed as a possible abstraction for the complex dynamical behaviors found in such large scale systems [6], [7]. Within this framework, the numerous sources of uncertainty including variations in heating load due to occupancy and equipment (often not directly measurable) can be potentially captured via a probabilistic model. Using such models, control designs for environment comfort can then be posed as safety control problems to satisfy the requirements of building codes, ANSI/ASHRAE standards [8], and owner specifications.

In literature, safety control problems are commonly studied within the domain of formal verification. Perfect information formulations, in which the state variables are assumed to be directly observed, have been studied extensively within the hybrid systems verification community (see for example [9]–[15]). In contrast, the case of partial information, with assumptions of incomplete or imprecise measurements of system state, has been seldomly considered. Due to the complexity of the problem, much of the work in this area has focused on systems with discrete state and observation spaces [16], [17], [18], or simple classes of hybrid systems to which the results in the discrete domain can be extended [19], [20]. In the case of deterministic hybrid systems with order preserving dynamics, safety control methods have also been proposed based upon set-valued estimates of the discrete or continuous state variables [21], [22].

In this work, we present an abstract dynamic programming solution to the partial information safety control problem for discrete time stochastic hybrid systems (DTSHS). This can be viewed as an extension of previous work on perfect information problems for such systems [14], [15], as well as a theoretical contribution towards the understanding of the role of estimation in a probabilistic safety control problem. In particular, by adapting the notion of a sufficient statistic [23], it is shown that the information needed for optimal safety control is a joint conditional distribution of the system state and a binary variable representing the safety of the state history. From this result, a dynamic programming procedure is derived for the abstract computation of the maximal safety probability and optimal control policy. While the development of practical control algorithms will require future investigations addressing computational complexity, these results provide insights into the structure of the optimal controller and estimator for safety specifications.

*This work has been supported in part by NSF under CPS:ActionWebs (CNS-931843), the European Commission under the MoVeS project, FP7-ICT-2009-5 257005 and Marie Curie grant MANTRAS PIRG-GA-2009-249295, and NWO under VENI grant 016.103.020.

[1]Jerry Ding and Claire Tomlin are with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, USA, {jding,tomlin}@eecs.berkeley.edu

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

II. Partially Observable Discrete Time Stochastic Hybrid Systems

The model for a partially observable discrete time stochastic hybrid system (POdtSHS) augments the perfect information stochastic hybrid system model proposed in [14] with an observation space and a stochastic observation model. For the rest of this paper, we denote the Borel σ-algebra of a topological space X by $\mathcal{B}(X)$.

Definition 1 (POdtSHS)**.** A partially observable discrete time stochastic hybrid system is a tuple $\mathcal{H} = (Q, n, U, Z, \nu_x, \nu_q, \nu_r, \zeta_0, \zeta)$, defined as follows.

- *Discrete state space* $Q := \{q_1, q_2, ..., q_m\}$, $m \in \mathbb{N}$;
- *Dimensions of continuous state space* $n : Q \to \mathbb{N}$: a map which assigns to each discrete state $q \in Q$ the dimension of the continuous state space $\mathbb{R}^{n(q)}$. The hybrid state space is given by $S := \bigcup_{q\in Q}\{q\} \times \mathbb{R}^{n(q)}$;
- *Control input space* U: a nonempty Borel space;
- *Observation space* Z: a nonempty Borel space;
- *Continuous state transition kernel* $\nu_x(dx'|(q, x), u)$: a Borel-measurable stochastic kernel which assigns to each $s = (q, x) \in S$ and $u \in U$ a probability measure on the Borel space $(\mathbb{R}^{n(q)}, \mathcal{B}(\mathbb{R}^{n(q)}))$;
- *Discrete state transition kernel* $\nu_q(q'|(q, x), u)$: a Borel-measurable stochastic kernel which assigns to each $s = (q, x) \in S$ and $u \in U$ a probability distribution over Q;
- *Reset transition kernel* $\nu_r(dx'|(q, x), u, q')$: a Borel-measurable stochastic kernel which assigns to each $s = (q, x) \in S$, $u \in U$, and $q' \in Q$ a probability measure on the Borel space $(\mathbb{R}^{n(q')}, \mathcal{B}(\mathbb{R}^{n(q')}))$;
- *Initial observation kernel* $\zeta_0(dz|s)$: a Borel-measurable stochastic kernel which assigns to each $s \in S$ a probability measure on the Borel space $(Z, \mathcal{B}(Z))$;
- *Observation kernel* $\zeta(\cdot|s, u)$: a Borel-measurable stochastic kernel which assigns to each $s \in S$ and $u \in U$ a probability measure on the Borel space $(Z, \mathcal{B}(Z))$.

Under a POdtSHS model $\mathcal{H}$, the available information at each time step k is the history of inputs and outputs $(z(0), u(0), ..., z(k-1), u(k-1), z(k))$, along with the prior distribution of the initial state $(q(0), x(0))$. For compactness of notation, we define as in [23] the information spaces

$$I_k := Z^{k+1} \times U^k, \ k = 0, 1, \ldots.$$

An element of I_k is referred to as the information vector at time step k. For the initial state distribution, we denote the set of probability measures on S by $\mathcal{P}(S)$. In the following, we define our optimization space to be the set of all randomized control policies with memory.

Definition 2. A policy π' for $\mathcal{H}$ is a sequence $\pi' = (\pi'_0, \pi'_1, ..., \pi'_{N-1})$ of universally measurable stochastic kernels $\pi'_k(du|p_0; i_k)$, which assigns to each initial state distribution $p_0 \in \mathcal{P}(S)$ and information vector $i_k \in I_k$ a probability measure on the Borel space $(U, \mathcal{B}(U))$. The set of such policies is denoted by Π'.

If for each k, initial state distribution $p_0 \in \mathcal{P}(S)$ and information vector $i_k \in I_k$ the stochastic kernel π'_k assigns probability mass one to some point in U, the policy π' is said to be *non-randomized*. The class of non-randomized policies for $\mathcal{H}$ is denoted as Π.

Following the procedure in [14], one can construct from ν_x, ν_q, and ν_r a hybrid state transition kernel $\nu(ds|s, u)$ on S given $S \times U$:

$$\nu((q', dx')|(q, x), u) = \tag{1}$$
$$\begin{cases} \nu_q(q|(q, x), u)\nu_x(dx'|(q, x), u), & \text{if } q' = q \\ \nu_q(q'|(q, x), u)\nu_r(dx'|(q, x), u, q'), & \text{if } q' \neq q. \end{cases}$$

For a given initial state distribution $p_0 \in \mathcal{P}(S)$ and policy $\pi' \in \Pi'$, the execution of the POdtSHS is as described in Algorithm 1.

Algorithm 1 POdtSHS Execution

Require: POdtSHS model $\mathcal{H}$, initial state distribution $p_0 \in \mathcal{P}(S)$, and control policy $\pi' \in \Pi'$.
Extract from S a value s_0 according to p_0;
Extract from Z a value z_0 according to $\zeta_0(\cdot|s_0)$;
Set $s(0) = s_0$ and $i_0 = z_0$;
for $k = 0$ to $N - 1$ **do**
  Extract from U a value u_k for $u(k)$ according to $\pi'_k(\cdot|p_0; i_k)$;
  Extract from S a value s_{k+1} for $s(k + 1)$ according to $\nu(\cdot|s_k, u_k)$;
  Extract from Z a value z_{k+1} for $z(k + 1)$ according to $\zeta(\cdot|s_{k+1}, u_k)$;
  Set $i_{k+1} = (i_k, u(k), z(k + 1))$;
end for
return Sample Path $\{(s_0, z_0, u_0, ..., s_N, z_N)\}$.

Now consider the sample space of state, observation, and control sequences over k time steps given by $\Omega_k := S^{k+1} \times Z^{k+1} \times U^k$. Then by Proposition 7.45 of [23], given $p_0 \in \mathcal{P}(S)$ and $\pi' \in \Pi'$, the stochastic kernels ν, ζ_0, and ζ induce a unique probability measure $P_k(\pi', p_0)$ on Ω_k, describing the probabilistic evolution of the closed-loop trajectory of system $\mathcal{H}$ under policy π'.

III. Problem Formulation

In this section, we give a formal definition of the probabilistic safety problem under partial information. Specifically, suppose that one is given a safe set $W \in \mathcal{B}(S)$ and a time horizon $[0, N]$. Then for a fixed initial state distribution $p_0 \in \mathcal{P}(S)$ and control policy $\pi' \in \Pi'$, the probability that the state trajectory $(s_0, s_1, ..., s_N)$ remains within the set W for every time instant $k = 0, 1, ..., N$ is given by

$$\begin{aligned} p_N^{\pi'}(p_0; W) :=& P_N(\pi', p_0)(\{(s_0, z_0, u_0, ..., s_N, z_N) : \\ & s_k \in W, \ \forall k \in [0, N]\}) \\ =& P_N(\pi', p_0)(W^{N+1} \times Z^{N+1} \times U^N). \end{aligned} \tag{2}$$

By Proposition 7.45 of [23], the safety probability above can be equivalently expressed as an expectation of a multiplicative payoff:

$$p_N^{\pi'}(p_0; W) = E_{p_0}^{\pi'}\left[\prod_{k=0}^{N} \mathbf{1}_W(s_k)\right], \tag{3}$$

where $E_{p_0}^{\pi'}$ denotes the expectation with respect to the probability measure $P_N(\pi', p_0)$ on the sample space Ω_N. Using this payoff, we define the partial information safety control problem as follows.

Problem 1. Given a POdtSHS $\mathcal{H}$, initial state distribution $p_0 \in \mathcal{P}(S)$, safe set $W \in \mathcal{B}(S)$, and time horizon N:

1) Compute the maximal probability of safety
$$p_N^*(p_0; W) := \sup_{\pi' \in \Pi'} p_N^{\pi'}(p_0; W);$$
2) Find an optimal policy $\pi^* \in \Pi'$, if it exists, such that $p_N^*(p_0; W) = p_N^{\pi^*}(p_0; W)$. Otherwise, for a choice of $\epsilon > 0$, find an ϵ-optimal policy $\pi_\epsilon^* \in \Pi'$ satisfying
$$p_N^{\pi_\epsilon^*}(p_0; W) \geq p_N^*(p_0; W) - \epsilon.$$

IV. Sufficient Statistic and Equivalent Perfect State Information Problem

A common approach to partial information stochastic optimal control problems is to transform the original problem into one of perfect information through the notion of sufficient statistic, which is, roughly speaking, an estimator which provides sufficient information for optimal control, with respect to the objective function of interest [23]. As will be shown in this section, a sufficient statistic for Problem 1 is an estimator which computes a joint conditional distribution of the current state and a binary random variable representing the safety of past state evolution.

A. Terminal Payoff Problem for an Augmented System

First, we will show that the partial information safety problem is equivalent to a terminal payoff problem when the state space is augmented with a binary variable. In particular, consider the random variables $h_k : \Omega_k \to \{0, 1\}$, $k = 0, 1, ..., N$, defined as:

$$h_0 := 1;\ h_k := \prod_{j=0}^{k-1} \mathbf{1}_W(s_j),\ k \geq 1. \tag{4}$$

These variables can be viewed as a binary state representing the safety of the state history up to time $k-1$. Now consider an augmented POdtSHS model with an expanded state space $\tilde{S} := \{0, 1\} \times S$, in which the state of the system at any time k is given by the pair (h_k, s_k). From (4),

$$h_{k+1} = \mathbf{1}_W(s_k) h_k,\ \forall k \geq 0,$$

which results in an augmented state transition kernel $\tilde{\nu}(d\tilde{s}|\tilde{s}, u)$ on $\tilde{S}$ given $\tilde{S} \times U$:

$$\tilde{\nu}((h_{k+1}, ds_{k+1})|(h_k, s_k), u_k) := \begin{cases} \nu(ds_{k+1}|s_k, u_k), & h_k = 0, h_{k+1} = 0 \\ 0, & h_k = 0, h_{k+1} = 1 \\ \mathbf{1}_{S\setminus W}(s_k)\nu(ds_{k+1}|s_k, u_k), & h_k = 1, h_{k+1} = 0 \\ \mathbf{1}_W(s_k)\nu(ds_{k+1}|s_k, u_k), & h_k = 1, h_{k+1} = 1. \end{cases} \tag{5}$$

For the augmented system, observation kernels $\tilde{\zeta}_0(dz|\tilde{s})$ and $\tilde{\zeta}(dz|\tilde{s}, u)$ can be simply defined as

$$\tilde{\zeta}_0(dz_k|h_k, s_k) := \zeta_0(dz_k|s_k), \tag{6}$$

$$\tilde{\zeta}(dz_k|h_k, s_k, u_{k-1}) := \zeta(dz_k|s_k, u_{k-1}). \tag{7}$$

With these definitions, we write the augmented POdtSHS model as $\tilde{\mathcal{H}} := (\tilde{S}, U, Z, \tilde{\nu}, \tilde{\zeta}_0, \tilde{\zeta})$.

Now consider a Borel-measurable function $\xi : \mathcal{P}(S) \to \mathcal{P}(\tilde{S})$ which takes an initial state distribution on S to an initial state distribution on $\tilde{S}$:

$$\xi(p_0)(h_0, ds_0) := \begin{cases} 0, & h_0 = 0 \\ p_0(ds_0), & h_0 = 1. \end{cases} \tag{8}$$

Clearly, ξ is one-to-one, which implies, by a technical result due to Kuratowski (see [24], p. 442, Corollary 2) that $\mathcal{P}(S)$ and $\xi(\mathcal{P}(S)) \subset \mathcal{P}(\tilde{S})$ are isomorphic Borel spaces, with the Borel isomorphism ξ.

We define the set of admissible control policies for an augmented POdtSHS model $\tilde{\mathcal{H}}$ as follows.

Definition 3. A policy $\tilde{\pi}'$ for $\tilde{\mathcal{H}}$ is a sequence $\tilde{\pi}' = (\tilde{\pi}'_0, \tilde{\pi}'_1, ..., \tilde{\pi}'_{N-1})$ of universally measurable stochastic kernels $\tilde{\pi}'_k(du_k|\xi(p_0); i_k)$, which assigns to each initial state distribution $\xi(p_0)$ and information vector $i_k \in I_k$ a probability measure on the Borel space $(U, \mathcal{B}(U))$. The set of such policies is denoted by $\tilde{\Pi}'$.

Given initial state distribution $\xi(p_0) \in \xi(\mathcal{P}(S))$ and policy $\tilde{\pi}' \in \tilde{\Pi}'$, the augmented stochastic kernels induce a probability measure $\tilde{P}_k(\tilde{\pi}', \xi(p_0))$ on the sample space $\tilde{\Omega}_k := \tilde{S}^{k+1} \times Z^{k+1} \times U^k$.

Now consider a terminal payoff problem for the augmented system, in which the objective function is the probability that the state trajectory reaches the set $\{1\} \times W$ at time N, as given by

$$\tilde{p}_N^{\tilde{\pi}'}(\xi(p_0); \{1\} \times W) := E_{\xi(p_0)}^{\tilde{\pi}'}\left[\mathbf{1}_{\{1\}\times W}(\tilde{s}_N)\right], \tag{9}$$

where $E_{p_0}^{\pi'}$ denotes the expectation with respect to the probability measure $\tilde{P}_N(\tilde{\pi}', \xi(p_0))$ on the sample space $\tilde{\Omega}_N$. The optimal payoff over the policy space $\tilde{\Pi}'$ is then

$$\tilde{p}_N^*(\xi(p_0); \{1\} \times W) := \sup_{\tilde{\pi}' \in \tilde{\Pi}'} \tilde{p}_N^{\tilde{\pi}'}(\xi(p_0); \{1\} \times W). \tag{10}$$

In the following, we establish the equivalence between Problem 1 and the terminal payoff problem (10).

Proposition 1. *Let $\mathcal{H}$ be a POdtSHS and $W \in \mathcal{B}(S)$ be a safe set. Let $\tilde{\mathcal{H}}$ be the corresponding augmented system. Then for every $p_0 \in \mathcal{P}(S)$, $N \in \mathbb{N}$, we have*

$$p_N^*(p_0; W) = \tilde{p}_N^*(\xi(p_0); \{1\} \times W).$$

Proof. Let $p_0 \in \mathcal{P}(S)$, $N \in \mathbb{N}$. Given that ξ is a Borel isomorphism, $\tilde{\Pi}'$ and Π' can be considered identical policy spaces via the identification $\pi'_k(du_k|p_0; i_k) = \tilde{\pi}'_k(du_k|\xi(p_0); i_k)$, for given policies $\pi' \in \Pi'$ and $\tilde{\pi}' \in \tilde{\Pi}'$. It is then sufficient to prove that, for every $\pi' \in \Pi'$, the following equality holds:

$$p_N^{\pi'}(p_0; W) = \tilde{p}_N^{\pi'}(\xi(p_0); \{1\} \times W).$$

Indeed, by the previous definitions,

$$\tilde{p}_N^{\pi'}(\xi(p_0);\{1\}\times W) = \int_{\tilde{\Omega}_N} \mathbf{1}_{\{1\}\times W}(\tilde{s}_N)\tilde{\zeta}(dz_N|\tilde{s}_N,u_{N-1})$$
$$\times \tilde{\nu}(d\tilde{s}_N|\tilde{s}_{N-1},u_{N-1})\pi'_{N-1}(du_{N-1}|p_0;i_{N-1})$$
$$\times \tilde{\zeta}(dz_{N-1}|\tilde{s}_{N-1},u_{N-2})\tilde{\nu}(d\tilde{s}_{N-1}|\tilde{s}_{N-2},u_{N-2})$$
$$\times \cdots \times \pi'_0(du_0|p_0;z_0)\tilde{\zeta}_0(dz_0|\tilde{s}_0)\xi(p_0)(d\tilde{s}_0)$$
$$= \int_{\tilde{S}\times S^N\times Z^{N+1}\times U^N} \prod_{k=1}^{N} \mathbf{1}_W(s_k)\zeta(dz_N|s_N,u_{N-1})$$
$$\times \nu(ds_N|s_{N-1},u_{N-1})\pi'_{N-1}(du_{N-1}|p_0;i_{N-1})$$
$$\times \cdots \times \pi'_0(du_0|p_0;z_0)\mathbf{1}_{\{1\}\times W}(\tilde{s}_0)\tilde{\zeta}_0(dz_0|\tilde{s}_0)\xi(p_0)(d\tilde{s}_0)$$
$$= \int_{\Omega_N} \prod_{k=0}^{N} \mathbf{1}_W(s_k)dP_N(\pi',p_0) = p_N^{\pi'}(p_0;W).$$

This completes the proof. □

B. Derivation of a Sufficient Statistic

We now proceed to derive a sufficient statistic with respect to the terminal payoff problem (10), using existing results for additive cost formulations of partial information problems. By Proposition 1, this in turn provides a sufficient statistic for the original safety control problem.

Intuitively, a sufficient statistic is an estimator which provides enough information to the controller for computation of expected future payoff. A formal definition is given below, as adapted from Definition 10.6 of [23].

Definition 4. A statistic for $\tilde{\mathcal{H}}$ is a sequence $(\eta_0,\eta_1,...\eta_{N-1})$ of Borel-measurable functions $\eta_k : \xi(\mathcal{P}(S))\times I_k \to B_k$, where $B_0,...,B_{N-1}$ are nonempty Borel spaces. A statistic $(\eta_0,\eta_1,...\eta_{N-1})$ for $\tilde{\mathcal{H}}$ is said to be sufficient for control if

1) For every $k = 0,1,...,N-1$, there exists a Borel-measurable stochastic kernel $\hat{\nu}(d\eta_{k+1}|\eta_k,u)$ on B_{k+1} given $B_k\times U$ such that for every $p_0\in\mathcal{P}(S)$, $\tilde{\pi}'\in\tilde{\Pi}'$, and $E_{k+1}\in\mathcal{B}(B_{k+1})$, the following identity holds
$$\tilde{P}_{k+1}(\tilde{\pi}',\xi(p_0))\{\eta_{k+1}(\xi(p_0);i_{k+1})\in E_{k+1}| \eta_k(\xi(p_0);i_k)=\eta,u_k=u\} = \hat{\nu}(E_{k+1}|\eta,u)$$
for $\tilde{P}_k(\tilde{\pi}',\xi(p_0))$ almost every (η,u).
2) There exists an upper semianalytic function $g_N : B_N\to[0,1]$ such that for every $p_0\in\mathcal{P}(S)$ and $\tilde{\pi}'\in\tilde{\Pi}'$, the following identity holds
$$E^{\tilde{\pi}'}_{\xi(p_0)}\left[\mathbf{1}_{\{1\}\times W}(\tilde{s}_N)|\eta_N(\xi(p_0);i_N)=\eta\right] = g_N(\eta)$$
for $\tilde{P}_N(\tilde{\pi}',\xi(p_0))$ almost every η.

Given the terminal payoff structure of (10), a sufficient statistic can be selected as a set of filtering equations which recursively update a conditional distribution on the augmented state space $\tilde{S}$. More precisely, consider a Borel-measurable mapping $\Psi:\mathcal{P}(\tilde{S})\times U\to\mathcal{P}(\tilde{S})$ defined as

$$\Psi(\tilde{p},u)(E) := \int_{\tilde{S}} \tilde{\nu}(E|\tilde{s},u)\tilde{p}(d\tilde{s}),\ \forall E\in\mathcal{B}(\tilde{S}). \tag{11}$$

Ψ can be viewed as the prediction step of a Bayesian filter. By Lemma 10.3 of [23], there also exist Borel-measurable stochastic kernels $\Phi_0(d\tilde{s}|\xi(p);z)$ on $\tilde{S}$ given $\xi(\mathcal{P}(S))\times Z$ and $\Phi(d\tilde{s}|\tilde{p};z,u)$ on $\tilde{S}$ given $\mathcal{P}(\tilde{S})\times Z\times U$ which satisfy

$$\int_{\tilde{S}'} \tilde{\zeta}_0(Z'|\tilde{s})\xi(p_0)(d\tilde{s}) = \tag{12}$$
$$\int_{\tilde{S}}\int_{Z'} \Phi_0(\tilde{S}'|\xi(p_0);z)\tilde{\zeta}_0(dz|\tilde{s})\xi(p_0)(d\tilde{s})$$
$$\int_{\tilde{S}'} \tilde{\zeta}(Z'|\tilde{s},u)\tilde{p}(d\tilde{s}) = \tag{13}$$
$$\int_{\tilde{S}}\int_{Z'} \Phi(\tilde{S}'|\tilde{p};z,u)\tilde{\zeta}(dz|\tilde{s},u)\tilde{p}(d\tilde{s})$$

for every Borel set $\tilde{S}'\in\mathcal{B}(\tilde{S})$, $Z'\in\mathcal{B}(Z)$, probability distribution $\xi(p_0)\in\xi(\mathcal{P}(S))$, $\tilde{p}\in\mathcal{P}(\tilde{S})$, and control input $u\in U$. Φ_0 and Φ can be viewed as, respectively, the initialization step and update step of a Bayesian filter.

For a given information vector $i_k\in I_k$ and initial state distribution $\xi(p_0)\in\xi(\mathcal{P}(S))$, define the stochastic kernels $\tilde{p}_k : \xi(\mathcal{P}(S))\times I_k\to\mathcal{P}(\tilde{S})$ recursively through the following filtering equations:

$$\tilde{p}_0(\xi(p_0);i_0) := \Phi_0(d\tilde{s}_0|\xi(p_0);z_0), \tag{14}$$
$$\tilde{p}_{k+1}(\xi(p_0);i_{k+1}) := \Phi(d\tilde{s}_{k+1}|\Psi(\tilde{p}_k,u_k);z_{k+1},u_k).$$

By Proposition 10.5 of [23], we have that the sequence $\{\tilde{p}_k(\xi(p_0);i_k)\}_{k=0}^{N-1}$ is a sufficient statistic for $\tilde{\mathcal{H}}$. In particular, there exists a transition kernel $\hat{\nu}(d\tilde{p}_{k+1}|\tilde{p}_k,u_k)$ describing the evolution of $\tilde{p}_k$, which we refer to as an *information state*. Moreover, one can define a terminal payoff with respect to the information state as

$$g_N(\tilde{p}_N) := \int_{\tilde{S}} \mathbf{1}_{\{1\}\times W}(\tilde{s}_N)\tilde{p}_N(d\tilde{s}_N). \tag{15}$$

C. Reduction to Perfect State Information Problem

We conclude this section by showing that Problem 1 is equivalent to a perfect information problem on the space of information states. In particular, consider a perfect state information model $\hat{\mathcal{H}}$ in which the state space is given by $\hat{S} := \mathcal{P}(\tilde{S})$, the action space is given by U, and the state transition kernel is given by $\hat{\nu}$. Now define the set of admissible control policies for $\hat{\mathcal{H}}$ as follows.

Definition 5. A policy $\hat{\pi}'$ for $\hat{\mathcal{H}}$ is a sequence $\hat{\pi} = (\hat{\pi}'_0,\hat{\pi}'_1,...,\hat{\pi}'_{N-1})$ of universally measurable stochastic kernels $\hat{\pi}'_k(du_k|\tilde{p}_0,u_0,...,\tilde{p}_{k-1},u_{k-1},\tilde{p}_k)$, which assigns to each sequence of controls and information states a probability measure on the Borel space $(U,\mathcal{B}(U))$. The set of such policies is denoted by $\hat{\Pi}'$.

We note that by Proposition 7.44 of [23], the policy space $\hat{\Pi}'$ can be viewed as a subset of $\tilde{\Pi}'$, and hence of Π'. In particular, it is the subset of control policies which depends upon the history of inputs and outputs only through the history of inputs and information states. If for each k, the stochastic kernel $\hat{\pi}'_k$ depends upon the history only through the current information state $\tilde{p}_k$, then the policy $\hat{\pi}'$ is said to be *Markov*. The class of non-randomized, Markov policies for $\hat{\mathcal{H}}$ is denoted as $\hat{\Pi}$.

Given an initial information state $\tilde{p}_0 \in \hat{S}$ and a policy $\hat{\pi}' \in \hat{\Pi}'$, the transition kernel $\hat{\nu}$ induce a probability measure $\hat{P}^{\hat{\pi}'}_{\tilde{p}_0}$ on the sample space $\hat{\Omega}_N := \hat{S}^{N+1} \times U^N$.

Now consider a terminal payoff function

$$J_{N,\hat{\pi}'}(\tilde{p}_0) := \int_{\hat{\Omega}_N} g_N(\tilde{p}_N) d\hat{P}^{\hat{\pi}'}_{\tilde{p}_0}. \tag{16}$$

and a perfect information optimal control problem for $\hat{\mathcal{H}}$

$$J^*_N := \sup_{\hat{\pi}' \in \hat{\Pi}'} J_{N,\hat{\pi}'}. \tag{17}$$

The next result establishes the equivalence between Problem 1 and the perfection information problem (17).

Proposition 2. *Let $\mathcal{H}$ be a POdtSHS and $W \in \mathcal{B}(S)$ be a safe set. Let $\hat{\mathcal{H}}$ be the corresponding perfect state information model. Define a function $\varphi : \mathcal{P}(S) \to \mathcal{P}(\hat{S})$ as*

$$\varphi(p_0)(E) := \int_S \zeta_0(\{z_0 | \tilde{p}_0(\xi(p_0); z_0) \in E\} | s_0) p_0(ds_0), \tag{18}$$

for every Borel set $E \in \mathcal{B}(\hat{S})$. Then we have

$$p^*_N(p_0; W) = \int_{\hat{S}} J^*_N(\tilde{p}_0) \varphi(p_0)(d\tilde{p}_0), \ \forall p_0 \in \mathcal{P}(S).$$

Furthermore, if $\hat{\pi}' \in \hat{\Pi}'$ is optimal, or ϵ-optimal for (17), then $\hat{\pi}'$ is also optimal, or ϵ-optimal for Problem 1.

The proof of this result proceeds by straightforward application of Proposition 10.3 of [23], combined with Proposition 1 of section IV-A, and is omitted for brevity.

V. Solution to Partial Information Safety Problem

In this section, we provide a dynamic programming solution to the partial information safety problem via its equivalent perfect information formulation defined in (17). Specifically, let $\mathcal{F}$ be the set of upper semianalytic functions from $\hat{S}$ to $[0, 1]$. Consider a dynamic programming operator $\mathcal{T} : \mathcal{F} \to \mathcal{F}$ defined as

$$\mathcal{T}(J)(\tilde{p}) := \sup_{u \in U} \int_{\hat{S}} J(\tilde{p}') \hat{\nu}(d\tilde{p}' | \tilde{p}, u), \ \tilde{p} \in \hat{S}. \tag{19}$$

Then the solution to (17) can be stated as follows.

Proposition 3. *Let $\hat{\mathcal{H}} = (\hat{S}, U, \hat{\nu})$ be a perfect state information model. Then*

1) *$J^*_N = \mathcal{T}^N(g_N)$;*
2) *For every $\epsilon > 0$, there exists an ϵ-optimal non-randomized Markov policy $\hat{\pi}^*_\epsilon \in \hat{\Pi}$ for (17).*

This result follows by standard dynamic programming results for perfect information problems (see for example Propositions 8.2, 8.3, and 10.1 of [23]). In other words, J^*_N can be computed through recursive applications of the dynamic programming operator $\mathcal{T}$, and that it is sufficient to consider the set of non-randomized Markov policies $\hat{\Pi}$ over the set of randomized policies $\hat{\Pi}'$. Thus, we arrive at our main result.

Theorem 1. *Let $\mathcal{H}$ be a POdtSHS and $W \in \mathcal{B}(S)$ be a safe set. Let $\hat{\mathcal{H}}$ be the corresponding perfect state information model. Define $g_N : \hat{S} \to [0, 1]$ as in (15) and $\varphi : \mathcal{P}(S) \to \mathcal{P}(\hat{S})$ as in (18). Then, for every $p_0 \in \mathcal{P}(S)$, we have*

1) *$p^*(p_0; W) = \int_{\hat{S}} \mathcal{T}^N(g_N)(\tilde{p}_0) \varphi(p_0)(d\tilde{p}_0)$;*
2) *For every $\epsilon > 0$, there exists an ϵ-optimal non-randomized policy $\pi^*_\epsilon \in \Pi$ for Problem 1 of the form*

$$\pi^*_{k,\epsilon}(p_0; i_k) = \hat{\pi}_{k,\epsilon}(\tilde{p}_k(\xi(p_0); i_k)), \ k = 0, 1, ..., N-1.$$

3) *If $\hat{\pi} = (\hat{\pi}_0, \hat{\pi}_1, ..., \hat{\pi}_{N-1}) \in \hat{\Pi}$ satisfies*

$$\hat{\pi}_k(\tilde{p}) \in \arg\sup_{u \in U} \int_{\hat{S}} J^*_{k+1 \to N}(\tilde{p}') \hat{\nu}(d\tilde{p}' | \tilde{p}, u), \ \tilde{p} \in \hat{S},$$

where $J^*_{k \to N} := \mathcal{T}^{N-k}(g_N)$, $k = 0, 1..., N-1$, then $\hat{\pi}$ is an optimal policy for Problem 1.

By this result, the maximal probability of safety $p^*(p_0; W)$ for a POdtSHS can be computed via a terminal payoff dynamic programming iteration. Furthermore, the ϵ-optimal policies can be found within the class of non-randomized policies which depends on the initial distribution p_0 and information vector i_k only through the information state $\tilde{p}_k$. This decouples the partial information safety problem into two subproblems:

1) Computing an ϵ-optimal control policy $\hat{\pi}^*_\epsilon$ using the dynamic programming recursion in Proposition 3;
2) Computing the information state $\tilde{p}_k$ through the filtering equations (14).

The first subproblem, which is the *control* aspect of the problem, can be performed in an offline setting, while the second subproblem, which is the *estimation* aspect of the problem, has to be performed in an online setting. It is important to remark that the numerical solutions to both of these problems require computational algorithms on the information state space $\hat{S}$. Given that $\tilde{p}_k$ is a probability distribution, which is in general infinite dimensional, rather than a hybrid state s_k, which is finite dimensional, this represents a significant growth in computational complexity over perfect information safety problems.

VI. Analytical Example

To illustrate some of the salient features of the partial information safety problem, as well as some of the main difficulties, we will discuss in this section a simple discrete state example. For such systems, the computational complexity can be concretely stated in terms of a class of well-known partial information optimal control problems.

In particular, consider a system $\mathcal{H}$ with the state space $S = \{q_1, q_2, q_3, q_4\}$, control input space $U = \{\sigma_L, \sigma_R\}$, and observation space $Z = \{o_L, o_R\}$. For $u = \sigma_L$, the state transition probability is defined as $\nu(q_j | q_i, \sigma_L) = 1$ if $i = j = 1$ or $j = i - 1$, $i = 2, 3, 4$, and $\nu(q_j | q_i, \sigma_L) = 0$ otherwise. For $u = \sigma_R$, the transition probability is defined as $\nu(q_j | q_i, \sigma_R) = 1$ if $i = j = 4$ or $j = i + 1$, $i = 1, 2, 3$, and $\nu(q_j | q_i, \sigma_R) = 0$ otherwise. The observation probability is characterized by $\zeta_0(o_L | s) = \zeta(o_L | s, u) = \alpha$ for $s = q_1, q_2$ and $\zeta_0(o_R | s) = \zeta(o_R | s, u) = \alpha$ for $s = q_3, q_4$, where $\alpha \in$

$[0.5, 1]$. For the safety control problem, the safe set is selected as $W = \{q_2, q_3\}$, and the initial state distribution is chosen to be $p_0 = (\frac{1-\beta}{2}, \frac{\beta}{2}, \frac{\beta}{2}, \frac{1-\beta}{2})$, $\beta \in [0, 1]$.

Systems of this type, with discrete state, input, and output spaces, are often referred to as Partially Observable Markov Decision Processes (POMDPs) [25], [26]. The information state $\tilde{p}_k$ for this example is a discrete distribution over the augmented state space $\tilde{S} = \{0, 1\} \times \{q_1, q_2, q_3, q_4\}$. More generally, one can show that the the safety problem for a POMDP is equivalent to a terminal payoff problem for an augmented POMDP with twice the number of discrete states. However, as shown in [27], the latter problem (i.e. the computation of an optimal policy with respect to the terminal payoff) is in general PSPACE-complete.

The simplicity of this particular example, however, allows the analytical calculation of an optimal solution. In particular, by applying Theorem 1, one can show that for any time horizon $N \geq 1$, the maximal probability of safety is given by $p_N^*(p_0; W) = \alpha\beta$, with a stationary optimal policy

$$\hat{\pi}_k^*(\tilde{p}) = \begin{cases} \sigma_R, & \tilde{p}_k(1, q_2) \geq \tilde{p}_k(1, q_3) \\ \sigma_L, & \text{otherwise.} \end{cases}$$

Further details can be found in section 5.7 of [28]. It is worth noting that the optimal policy requires knowledge of an augmented state distribution over the space $\tilde{S}$, rather than simply a conditional state distribution over S.

VII. Future Work

There are several possible future research directions for addressing the computational challenges of a partial information safety problem. First, one may investigate finite dimensional representations or approximations of the information state for subclasses of stochastic hybrid systems, in order to allow the computation of a control policy on a finite dimensional space. Another direction is to find methods for computing optimal control policies with respect to specific choices of estimation schemes. The resulting safety performance can then be compared to decide on an appropriate design. Finally, in the case that the measurement uncertainty is bounded, it may also be possible to take a robust control approach by treating the uncertainty as a source of disturbance, albeit at some cost of conservativeness.

References

[1] 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.

[2] J. Lygeros, D. Godbole, and S. Sastry, "Verified hybrid controllers for automated vehicles," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 522–539, 1998.

[3] E. Frazzoli, M. Dahleh, and E. Feron, "Robust hybrid control for autonomous vehicle motion planning," in *Proceedings of the 39th IEEE Conference on Decision and Control*, vol. 1, December 2000, pp. 821–826.

[4] 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. Lynch and B. Krogh, Eds. Springer Berlin / Heidelberg, 2000, vol. 1790, pp. 160–173.

[5] C. G. Cassandras and J. Lygeros, Eds., *Stochastic Hybrid Systems*. Boca Raton, FL: CRC Press, 2006.

[6] A. Aswani, N. Master, J. Taneja, V. Smith, A. Krioukov, D. Culler, and C. Tomlin, "Identifying models of HVAC systems using semiparametric regression," in *Proceedings of the American Control Conference*, June 2012, pp. 3675–3680.

[7] "MoVeS website," http://www.movesproject.eu.

[8] "Thermal Environmental Conditions for Human Occupancy," *ANSI/ASHRAE Standard 55-2010*, 2010.

[9] E. Asarin, O. Bournez, T. Dang, O. Maler, and A. Pnueli, "Effective synthesis of switching controllers for linear systems," *Proceedings of the IEEE*, vol. 88, no. 7, pp. 1011–1025, 2000.

[10] A. Kurzhanski and P. Varaiya, "Ellipsoidal techniques for reachability analysis," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, N. Lynch and B. Krogh, Eds. Springer Berlin / Heidelberg, 2000, vol. 1790, pp. 202–214.

[11] J.-P. Aubin, J. Lygeros, M. Quincampoix, S. Sastry, and N. Seube, "Impulse differential inclusions: a viability approach to hybrid systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 1, pp. 2–20, 2002.

[12] X. 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. Hespanha and A. Tiwari, Eds. Springer Berlin / Heidelberg, 2006, vol. 3927, pp. 377–391.

[13] A. Girard and C. Le Guernic, "Zonotope/hyperplane intersection for hybrid systems reachability analysis," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, M. Egerstedt and B. Mishra, Eds. Springer Berlin / Heidelberg, 2008, vol. 4981, pp. 215–228.

[14] 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, 2008.

[15] 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.

[16] R. Cieslak, C. Desclaux, A. Fawaz, and P. Varaiya, "Supervisory control of discrete-event processes with partial observations," *IEEE Transactions on Automatic Control*, vol. 33, no. 3, pp. 249–260, 1988.

[17] F. Lin and W. Wonham, "On observability of discrete-event systems," *Information Sciences*, vol. 44, no. 3, pp. 173–198, 1988.

[18] K. Chatterjee, L. Doyen, T. Henzinger, and J.-F. Raskin, "Algorithms for omega-regular games with imperfect information," in *Computer Science Logic*, ser. Lecture Notes in Computer Science, Z. Ésik, Ed. Springer Berlin / Heidelberg, 2006, vol. 4207, pp. 287–302.

[19] H. Wong-Toi, "The synthesis of controllers for linear hybrid automata," in *Proceedings of the 36th IEEE Conference on Decision and Control*, vol. 5, December 1997, pp. 4607–4612.

[20] M. De Wulf, L. Doyen, and J.-F. Raskin, "A lattice theory for solving games of imperfect information," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, J. Hespanha and A. Tiwari, Eds. Springer Berlin / Heidelberg, 2006, vol. 3927, pp. 153–168.

[21] D. Del Vecchio, M. Malisoff, and R. Verma, "A separation principle for a class of hybrid automata on a partial order," in *Proceedings of the American Control Conference*, June 2009, pp. 3638–3643.

[22] R. Verma and D. Del Vecchio, "Safety control of hidden mode hybrid systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 62–77, 2012.

[23] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete Time Case*. New York, NY: Academic Press, 1978.

[24] K. Kuratowski and A. Mostowski, *Set Theory: With an Introduction to Descriptive Set Theory*. Amsterdam: North-Holland, 1976.

[25] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach, 2nd Ed.* Englewood Cliffs, NJ: Prentice Hall, 2002.

[26] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA: MIT Press, 2005.

[27] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of Markov Decision Processes," *Mathematics of Operations Research*, vol. 12, no. 3, pp. 441–450, 1987.

[28] J. Ding, "Methods for reachability-based hybrid controller design," Ph.D. dissertation, University of California, Berkeley, 2012. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-80.html