Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

ThC11.6

A Statistical Inference Method for the Stochastic Reachability Analysis

Manuela L. Bujorianu

Abstract— The main contribution of this paper is the characterization of reachability problem associated to stochastic hybrid systems in terms of imprecise probabilities. This provides the connection between reachability problem and Bayesian statistics. Using generalised Bayesian statistical inference, a new concept of conditional reach set probabilities is defined. Then possible algorithms to compute the reach set probabilities are derived.

Keywords: reachability problem, stochastic hybrid automata, Markov process, Choquet capacity, Bayesian inference.

I. INTRODUCTION

Within the Hybridge project [18] we have modeled the Air Traffic Control (ATC) using the stochastic hybrid system paradigm. The centralised ATC is in complete control of the air traffic and ultimately responsible for the aircraft safety. The main objective of ATC is to maintain safe separation between aircraft. To improve the ATC performance, research effort has spent in the last decade to create tools for Conflict Detection and Conflict Resolution. Conflict Detection means the evaluation of the possibility of future conflict starting from the current state of the airspace and taking into account the uncertainty in the future position of the aircraft while they follow given nominal paths. A possible approach is to determine the conflict probability over a certain time horizon, i.e. the probability that two aircraft come closer than a minimum allowed distance. If this probability can be computed, an alert can be issued when it exceeds a certain threshold. Conflict Resolution means to calculate suitable maneuvers to avoid a predicted conflict.

In the context of stochastic hybrid systems, the computation of the conflict probability reduces to a *reachability problem*: *computing the probability that the stochastic hybrid process modelling the aircraft motion reaches an unsafe part of the state space* (where two aircraft come closer than the minimum allowed distance).

In a probabilistic framework, the stochastic reachability problem consists of determining the probability that the system trajectories enter some prespecified set starting from a certain set of initial conditions with a given probability distribution.

It has been proven [4], [5] that for the reach set probabilities it is hard to find analytical formulas. At least, one might get upper bounds of these, but again these bounds can not be easily computed. Then, it seems to be the case to set up a statistics framework which will allow us to find suitable algorithms to compute these probabilities. In this paper we characterize the stochastic reachability in terms

of Bayesian statistics. This will permit us to employ statistical inference rules to compute the reach set probabilities which appear in the formulation of stochastic reachability problem. To achieve this task, the stochastic reachability concept will be expressed in terms of imprecise probabilities. Here, imprecision in probability assessments is modeled through convex sets of probability measures. This is possible since we consider only stochastic hybrid systems whose realizations are Markov processes with nice properties. Then the reach set probabilities define a Choquet capacity. The later concept is widely used in decision theory [11], [25], robust Bayesian inference [16] and is closely related to other concepts modelling different kinds of probability sets as: lower probabilities [12], belief functions [11], lower envelopes [16], lower expectations, lower previsions [28].

The paper is organized as follows. Section II presents the necessary background on capacities. In section III, we give a general concept of stochastic hybrid automaton and define the corresponding stochastic reachability problem. Then the later problem is characterized in terms of Choquet capacities. Different algorithms used in statistical inference (for the computation of conditional upper expectations) are proposed to calculate the reach set probabilities. The paper ends with some conclusions.

II. PRELIMINARIES

Decision theory starts from the states, acts and utilities that have to be specified by the acting agent. Making a decision is deciding which possible act to follow. Bayesian theory has been very successful in this regard as a prescription for what a rational agent should do [23]. The Bayesian framework essentially says that:

• Given the states of nature θ_i, there is a single probability distribution $p(\theta)$ that summarizes the beliefs of the agent about which θ_i obtains.

• An act with high expected utility is preferred to an act with lower expected utility.

The Bayesian framework is derived from a number of axioms that are supposed to apply to decision making.

The next step is to start with similar, but more general set of axioms and generate a convex set of probability distributions, called *credal set* [21]. In this context, Bayesian theory is a particular case in which we assume that the agent always has a single distribution (the convex set of distributions has a single member).

The theory of sets of probabilities advocates that a rational agent chooses an act based on expected loss considerations. Expected loss are defined by:

M.L. Bujorianu is with Faculty of Computer Science, University of Twente, 7500 AE Enschede, The Netherlands, mlb@cs.stir.ac.uk

1. A loss function, which translates the preferences of the agent. Some people use the term utility for the reverse of loss, i.e. loss with a minus sign. But loss and utility are essentially the same thing.

2. A set of probability distributions called credal set. Usually a credal set is assumed to be a convex set of probability distributions. A credal set conveys the beliefs of an agent about the possible states of the world.

Usually, a theory of sets of probabilities is a normative theory of decision making. So the purpose is to explain how an agent should make decisions. The agent starts with a prior credal set, uses a likelihood credal set and reaches a posterior credal set. Then the agent picks an option that minimizes expected loss.

A number of theories of inference advocate close convex sets of probability measures as an accurate representation for imprecise belief [21], [19]. Several other theories employ special types of convex sets of probability measures, for example the theory of lower probability [12] and the of inner/outer measures [27]. Others are based on Choquet capacities, monotone (convex) capacities, 2-monotone capacities, infinitely monotone capacities (belief functions) [16]. The theory of coherent lower previsions (lower expectations) put forward by Walley is an example of a complete theory of inference that can be viewed as a theory of sets of probability measures [28]. There are also theories of inference that add imprecision in utility judgements to the modelling process [26]. Theories are normative because they only offer some sensible guidelines; they are not supposed to be a description of how real agents work.

In the following we present the concept of capacity and the Bayes theorem for capacities.

A. Capacities

For every space X and algebra $\mathcal{A}$ of subsets of X a set-function $c : \mathcal{A} \to [0, 1]$ is called a *normalized capacity* if it satisfies the following: (i) $c(\emptyset) = 0$, $c(X) = 1$, (ii) $\forall A, B \in \mathcal{A}$: $A \subset B \Rightarrow c(A) \leq c(B)$.

A capacity is called *convex (or supermodular)* if in addition to (i)-(ii) it satisfies the property
(iii) $\forall A, B \in \mathcal{A}$: $c(A \cup B) \geq c(A) + c(B) - c(A \cap B)$.

A special type of convex capacities are the *belief functions* presented and discussed by Dempster [11] and Shafer [24]. A capacity becomes a probability if (iii) holds everywhere with equality, i.e. it is additive. If a capacity satisfies the inverse inequality in (iii) then it is called *submodular or strongly subadditive*.

If X is a topological space, c is a Choquet capacity if $c(F_n) \downarrow c(F)$ for each sequence of closed sets $F_n \downarrow F$.

Since we allow the possibility that c is not additive, we can not use the integral in the Lebesgue sense to integrate w.r.t. c. The notion of integral we will use is due originally to Choquet [7] and it was independently rediscovered and extended by Schmeidler [25]. If $f : X \to \mathbb{R}$ is a bounded $\mathcal{A}$-measurable function and c is any capacity on X, the Choquet integral of f w.r.t. c is defined as follows

$$
\begin{aligned}
c(f) &= \int_X f(x)dc(x) = \overline{c}(f) + \underline{c}(f) \\
\overline{c}(f) &= \int_0^\infty c(\{x \in X | f(x) \geq \alpha\})d\alpha \\
\underline{c}(f) &= \int_{-\infty}^0 [c(\{x \in X | f(x) \geq \alpha\}) - 1]d\alpha
\end{aligned}
$$

where the integrals are taken in the sense of Riemann. Then $c(f)$ can be thought of as an expected utility without additivity [25].

B. Bayes Theorem for Capacities

Let $(X, \mathcal{B})$ a Polish space, i.e. the topology for X is complete, separable, and metrizable, and $\mathcal{B}$ is the Borel σ-algebra of X. Denote by $\mathcal{B}^b(X)$ the set of all bounded, nonnegative, real-valued, measurable functions defined on X. Let $(P_x | x \in X)$ be a set of probability measures on a sample space $(\Omega, \mathcal{F})$. Assume that each P_x has a density $p(x|\omega)$ w.r.t. some σ-finite measure and let $L(x) = p(x|\omega)$ be the likelihood function for x having observed $\omega \in \Omega$. We assume that $L \in \mathcal{B}^b(X)$.

Let $\mathcal{P}$ be a nonempty set of prior probabilities on $\mathcal{B}$ and define the upper and lower prior probability functions by

$$
\overline{\mu}(A) = \sup_{\mu \in \mathcal{P}} \mu(A), \quad \underline{\mu}(A) = \inf_{\mu \in \mathcal{P}} \mu(A) \tag{1}
$$

Clearly, $\overline{\mu}(A) = 1 - \underline{\mu}(X \setminus A)$, then it is enough to study only $\overline{\mu}$. Suppose that $\mathcal{P}$ is convex. $\overline{\mu}$ can be thought of as the capacity c generated by the family $\mathcal{P}$ (the upper envelope of $\mathcal{P}$). For each $f \in \mathcal{B}^b(X)$, define the upper expectation of f by

$$
\overline{E}(f) = \sup_{\mu \in \mathcal{P}} \mu(f) \tag{2}
$$

where $\mu(f) = \int f(x)\mu(dx)$. Analogously, the lower expectation $\underline{E}(f)$ can be defined.

In a similar way, one can define the upper Choquet integral of f defined by

$$
c_u(f) = \overline{c}(f) = \int_0^\infty \overline{\mu}(\{x \in X | f(x) \geq \alpha\})d\alpha
$$

and the lower Choquet integral of f by

$$
c_l(f) = \int_0^\infty \underline{\mu}(\{x \in X | f(x) \geq \alpha\})d\alpha
$$

We say that $\mathcal{P}$ is closed w.r.t. majorization, or is m-close, if $\mu \leq \overline{\mu}$ implies that $\mu \in \mathcal{P}$.

For a prior probability $\mu \in \mathcal{P}$, for which $\mu(L) > 0$, the posterior probability of a subset A, after observing ω and applying Bayes' theorem, may be expressed as

$$
\mu(A|\omega) = \frac{\mu(L_A)}{\mu(L_A) + \mu(L_{A^c})}
$$

where $L_A(x) = L(x)I_A(x)$ and I_A is the indicator function of A.

We denote $\mathcal{P}(\cdot|\omega) = \{\mu(\cdot|\omega) | \mu \in \mathcal{P} \text{ with } \mu(L) > 0\}$. Next we give the Bayes theorem for Choquet capacities as a particular case of the main result of [29].

Theorem 1: Let $\mathcal{P}$ be a nonempty m-close set of prior probabilities on $\mathcal{B}$ and let $\mathcal{P}(\cdot|\omega)$ be the corresponding class of posterior probabilities. If $\mathcal{P}$ generates a submodular Choquet capacity c, then for each $A \in \mathcal{B}$, the capacity generated as the upper envelope of $\mathcal{P}(\cdot|\omega)$ can be computed as follows

$$
\begin{aligned}
c(A|\omega) &= \frac{\overline{E}(L_A)}{\overline{E}(L_A) + \underline{E}(L_{A^c})} \\
&= \frac{c_u(L_A)}{c_u(L_A) + c_l(L_{A^c})}
\end{aligned}
$$

when the ratios are well defined.

Remark 1: If in Th.1, the capacity c is not a Choquet capacity, then both equals should be replaced by '$\leq$'. Then, in this case, the Bayes theorem for capacities gives upper bounds of the conditional capacity $c(\cdot|\omega)$.

III. STATISTICAL INFERENCE FOR STOCHASTIC REACHABILITY ANALYSIS

In this section we give the concept of stochastic hybrid automaton (SHA), in a general setting. In most of the cases of stochastic hybrid automata studied in the literature [4], [5], [22], [2], the realisation of a stochastic hybrid automaton is a Markov process with nice properties. We briefly present the construction of the sample probability space of an SHA realisation and of the Choquet capacity associated. To the reader less interested in the rigour of the mathematical foundations of this method we propose a reading path that skips the subsection III-C and uses only the formula (8). Then we link the concept of stochastic reachability defined in [4] with the concept of capacity. This connection is of crucial importance for the purpose of computing reach set probabilities because it allows one to make use of Bayesian inference techniques.

A. Stochastic Hybrid Systems

Formally, a stochastic hybrid automaton (SHA) is defined as a tuple $H = (Q, \mathcal{X}, F, R, \lambda)$
- Q is a countable set of discrete variables;
- $\mathcal{X} : Q \to \mathbb{R}^{d(\cdot)}$ maps each $q \in Q$ into a mode (an open subset) X^q of $\mathbb{R}^{d(q)}$, where $d(q)$ is the Euclidean dimension of the corresponding mode;
- $F : Q \to 2^{\mathcal{F}_{SDE}}$ specifies the continuous evolution of the automaton in terms of stochastic differential equations over the continuous state x^q for each mode;
- $R = (R^q)_{q \in Q}$ a family of stochastic kernels R^q : $\overline{X}^q \times \underset{j \in Q \setminus \{q\}}{\cup} \mathcal{B}(X^j) \to [0, 1]$;
- $\lambda : \underset{j \in Q}{\cup} \overline{X}^j \to \mathbb{R}^+$ is a transition rate function which gives the probability distributions of the jumping times.

The executions of an SHA can be described as follows: start with an initial point $x_0 \in X^q$, follow a solution of the SDE associated to X^q, jump when this trajectory hits the boundary or according with the transition rate λ (the jump time is the minimum of the boundary hitting time and the time which is exponentially distributed with the transition rate λ). Under standard assumptions, for each

initial condition $x \in \underset{j \in Q}{\cup} X^j$, the possible trajectories, which start from x, form a stochastic process. Moreover, for all initial conditions x, the realizations of an SHA make up a family of Markov processes, which can be thought a Markov process in a general approaching (see [10] for more explanations).

Let us consider $M = (\Omega, \mathcal{F}, \mathcal{F}_t, x_t, P_x)$, the realization of a stochastic hybrid system H, as a family of Markov processes with the state space $(X, \mathcal{B})$ (X is the union of modes and $\mathcal{B}$ is its Borel σ-algebra). That means
- $\mathcal{F}_t^0$ denotes the *natural filtration*, i.e. $\mathcal{F}_t^0 = \sigma\{x_t, s \leq t\}$ for all $t \in [0, \infty]$.
- $P_x : (\Omega, \mathcal{F}) \to [0, 1]$ is a probability measure (called *Wiener probability*) such that $P_x(x_t \in E)$ is $\mathcal{B}$-measurable in $x \in X$ for each $t \geq 0$ and $E \in \mathcal{B}$.
- If μ is a probability on $(X, \mathcal{B})$ then we can define

$$
P_\mu(\Lambda) = \int_X P_x(\Lambda)\mu(dx), \Lambda \in \mathcal{F}.
$$

We then denote by $\mathcal{F}_t$ the completion of $\mathcal{F}_t^0$ w.r.t. all P_μ.
- $(\Omega, \mathcal{F}, P_x)$ will be the sample probability space of a Markov process of this family with the initial probability distribution $P_x(x_0 = x) = 1$.

B. Stochastic Reachability Problem

Let us consider $M = (\Omega, \mathcal{F}, \mathcal{F}_t, x_t, P_x)$, the realization of a stochastic hybrid automaton H. To address the reachability problem assume that we have a given set $A \in \mathcal{B}(X)$ and a time horizon $T > 0$. Let us to define:

$$
\begin{aligned}
Reach_T(A) &= \{\omega \in \Omega \mid \exists t \in [0, T] : x_t(\omega) \in A\} \\
Reach_\infty(A) &= \{\omega \in \Omega \mid \exists t \geq 0 : x_t(\omega) \in A\}. \quad (3)
\end{aligned}
$$

These two sets are the sets of trajectories of H, which reach the set A (the flow that enters A) in the interval of time $[0, T]$ or $[0, \infty)$. The reachability problem consists of determining the probability measures of such sets

$$
P_x(Reach_T(A)) \text{ and } P_x(Reach_\infty(A)). \quad (4)
$$

The reachability problem is well-defined, i.e. $Reach_T(A)$ and $Reach_\infty(A)$ are indeed measurable sets [4], [5].

The probabilities (4) can be expresses as

$$
E_x \sup_{t \in [0,T]} I_A(x_t) \text{ and } E_x \sup_{t \geq 0} I_A(x_t) \quad (5)
$$

where E_x is the expectation w.r.t. P_x. The computation of the reach set probabilities (4) is equivalent to the computation of the following probabilities

$$
P_x(T_A < T) \text{ or } P_x(T_A < \infty) \quad (6)
$$

for each initial state $x \in X$, where T_A is the first hitting time of A, i.e.

$$
T_A = \inf\{t > 0 | x_t \in A\}.
$$

Though the reach set probability expressions (5) are given as expectations of some measurable functions defined along the trajectories their computations is not easy. The main

difficulty is that the integrands do not have any monotonicity, they only 'account' the trajectories which enter A in the time interval $[0, T]$ or $[0, \infty)$. As well, it is hard to give an analytical expression of these integrands.

Moreover, the expressions (5) or (6) of reach set probabilities show that these measures as functions of the target A are not additive. It can be easily checked that

$$P_x(Reach_T(A \cup B)) \leq P_x(Reach_T(A)) + P_x(Reach_T(B))$$

and a similar inequality for $P_x(Reach_\infty(A))$. Then, the idea to 'approximate' these with set of additive measures seems to be natural. This will be mathematically described in the next subsections.

C. Analysis of SHA Realisation

Hypothesis

We assume without loss of generality that the realization of an SHA $M = (\Omega, \mathcal{F}, \mathcal{F}_t, x_t, P_x)$ is a Borel right Markov process with the state space $(X, \mathcal{B})$. This is the case with the most of models for stochastic hybrid systems existing in the literature [22]. This means that

1. M is a strong Markov process with paths that are right-continuous with left limits (cadlag property),
2. X is a separable metric space homeomorphic to a Borel subset of some compact metric space, equipped with Borel σ-algebra $\mathcal{B}(X)$ or shortly $\mathcal{B}$. That means X is a Lusin state space and it will be equipped with a σ-finite measure m.
3. the operator semigroup of M (given by $P_t f(x) = E_x f(x_t)$, where E_x is the expectation w.r.t. P_x) maps $\mathcal{B}^b$

(the lattice of bounded real measurable functions defined on X) into itself.

We assume also that M is *transient*. This means that there exists a strictly positive Borel function q such that Uq is bounded (where $Uf = \int_0^\infty P_t f dt$ is the *kernel operator*). More, we suppose that $\sup_{x \in X} U1(x) < \infty$. The transience of M means that any process trajectory which will visit a Borel set of the state space it will leave it after a finite time.

Representation of the sample probability space

Let Δ be the cemetery point for X, which is an adjoined point to X, $X_\Delta = X \cup \{\Delta\}$. The existence of Δ is assumed in order to have a probabilistic interpretation of $P_x(x_t \in X) < 1$, i.e. at some 'termination time' $\zeta(\omega)$ when the process M escapes to and is trapped at Δ. One can take the sample space Ω for M to be the set of all paths $(0, \infty) \ni t \mapsto \omega_t \in X_\Delta$ such that (i) $t \mapsto \omega_t$ is X-valued and cadlag on $(0, \zeta(\omega))$ where $\zeta(\omega) := \inf\{s > 0 | \omega(s) = \Delta\}$, (ii) $\omega_t = \Delta$ for all $t \geq \zeta(\omega)$, and (iii) $\zeta(\omega) < \infty$. In this way, M is realized as the coordinate process on Ω: $x_t(\omega) = \omega_t$, $t > 0$. We complete the definition of M by declaring $x_0(\omega) = \lim_{t \searrow 0} \omega_t$, $t > 0$. Then, the trajectories of the process M can be thought of as elements in the sample probability space, that means each $\omega \in \Omega$ is a process evolution, i.e. $\omega = (\omega_t)_{t \geq 0}$.

The process (x_t) can be viewed as the signal process and (ω_t) as the observable process.

The global probability measure $\mathbb{P}$

Because of transience condition, the measure m is purely excessive [15]:

$$\lim_{t \to \infty} (m < P_t >)(A) = 0, \forall A \in \mathcal{B} \text{ with } m(A) < \infty,$$

where $(m < P_t >)(A) = \int p_t(x, A) m(dx)$ and

$$p_t(x, A) = P_t(I_A)(x) = P_x(x_t \in A). \tag{7}$$

Consequently, there is a unique entrance law $(\mu_t)_{t>0}$ (a family of σ-finite measures on $(X, \mathcal{B})$ with $\mu_t < P_s >= \mu_{t+s}$ for all $t, s > 0$) such that

$$m(A) = \int_0^\infty \mu_t(A) dt, \forall A \in \mathcal{B}.$$

See, for example [15], for more details. Then there is a σ-finite measure $\mathbb{P}$ on $(\Omega, \mathcal{F}_t^0)$ under which the coordinate process $(x_t)_{t>0}$ is Markovian with transition semigroup $(P_t)_{t \geq 0}$ and one-dimensional distributions

$$\mathbb{P}(x_t \in A) = \mu_t(A), \forall A \in \mathcal{B}, t > 0.$$

The contruction of the capacity

The *capacity* associated to M is defined as follows (see [14] and the references therein): for all $B \in \mathcal{B}$

$$Cap_M(B) = \mathbb{P}(T_B < \infty), \tag{8}$$

where T_B is the first hitting time of B.

The initial definition of this notion gives the capacity Cap_M as an upper envelope of a non-empty class of probability measures on $\mathcal{B}$. It can be shown that this capacity is *monotone increasing, submodular, countably subadditive* [14]. Then its conjugate Cap_M^* [25], defined by $Cap_M^*(B) = 1 - Cap_M(B^c)$ is a *belief function* in sense of [24] (here, $B^c = X - B$). Beliefs about the evolution process (ω_t) conform to a time-homogeneous Markov structure. In standard models, this would involve a stochastic kernel giving conditional probabilities. We assume that beliefs conditional on ω_t are too vague to be represented by a probability measure and are represented instead by a family of probability measures whose lower envelope is Cap_M^*.

Let us consider $\mathcal{P}_M$ the family of all probability measures on $(X, \mathcal{B})$ dominated by Cap_M. Then, it is known from [16],

$$Cap_M(A) = \overline{\mu}_M(A) = \sup_{\mu \in \mathcal{P}_M} \mu(A). \tag{9}$$

D. Generalized Bayes Rule for Stochastic Reachability

The construction of the 'global' probability measure $\mathbb{P}$ defined on the sample probability space of M allows us to replace (6) by

$$\mathbb{P}(T_A < T) \text{ or } \mathbb{P}(T_A < \infty).$$

where $A \in \mathcal{B}$ is a target set and $T > 0$. In this way, the reachability problem is related with the computation of the capacities associated to the processes M_T and M, where M_T

is the process M "killed" after the time T (see [10] for the details about the killed process), i.e.

$$\mathbb{P}(Reach_T(A)) = Cap_{M_T}(A)$$
$$\mathbb{P}(Reach_\infty(A)) = Cap_M(A)$$

Then the computation of the reach set probabilities associated to A reduces to the calculating of the capacity of A. Since Cap_M is a submodular Choquet capacity, the Theorem 1 (Bayes theorem for capacities) is applicable under some mild conditions as follows.

Let suppose that
• Each P_x has a density $p(x|\omega)$ w.r.t. some σ-finite measure on $(\Omega, \mathcal{F})$. Then the likelihood function $L(x) = p(x|\omega)$ is constructed as in the subsection II-B. Considering the time horizon $T > 0$, we will also consider the likelihood function $L_T(x) = p(\omega_T, x)$ for x having observed ω_T (the evolution process until T).
• $\mathcal{P} = \mathcal{P}_M$ (resp. $\mathcal{P}_T = \mathcal{P}_{M_T}$) be a $m-$close convex family of probability measures (i.e. a credal set) on $(X, \mathcal{B})$ dominated by Cap_M (resp. Cap_{M_T}).

Following [13], for each $\mu \in \mathcal{P}$ there exists a measurable function f_μ, with $0 \leq f_\mu \leq 1$ such that

$$\mu(f_\mu) = Cap_M(f_\mu). \tag{10}$$

and for each bounded measurable function f there exists $\mu_f \in \mathcal{P}$ such that

$$\mu_f(f) = Cap_M(f). \tag{11}$$

Similar results are true for any $\mu \in \mathcal{P}_T$.

Applying Th.1 to Cap_M (resp. Cap_{M_T}), we get the following estimations

$$Cap_M(A|\omega) = \frac{\overline{E}(L_A)}{\overline{E}(L_A) + \underline{E}(L_{A^c})} \tag{12}$$

$$Cap_M(A|\omega_T) = \frac{\overline{E}_T(L_{T,A})}{\overline{E}_T(L_{T,A}) + \underline{E}_T(L_{T,A^c})} \tag{13}$$

where $\overline{E}_T$ is the upper expectation defined w.r.t. $\mathcal{P}_T$. Note that

$$Cap_M(L_A) = \int p(x|\omega) I_A(x) \mu_{L_A}(dx)$$

where μ_{L_A} is defined as in (11).

The formulas (12) and (13) give the idea to introduce expressions for *conditional reach set probabilities* (i.e. the probabilities to reach A having observed the trajectory ω_T or ω)

$$\mathbb{P}(Reach_T(A|\omega_T)) = Cap_{M_T}(A|\omega_T)$$
$$\mathbb{P}(Reach_\infty(A|\omega)) = Cap_M(A|\omega) \tag{14}$$

In the following we consider only the reachability problem in infinite time horizon and respectively only the capacity Cap_M. The case of the reachability problem with finite time horizon can be treated in a similar way taking the process M_T. The 'conditional capacity' formula (12) can be extended

for the case we have a set of trajectories E observed (after a 'learning process'). We get

$$\mathbb{P}(Reach_\infty(A|E)) = Cap_M(A|E) \tag{15}$$
$$= \frac{\overline{E}(L_{A,E})}{\overline{E}(L_{A,E}) + \underline{E}(L_{A^c,E})}$$
$$= \frac{Cap_M^u(L_{A,E})}{Cap_M^u(L_{A,E}) + Cap_M^l(L_{A^c,E})}$$

where $L_{A,E}(x) = p(x|\omega) I_A(x) I_E(\omega)$ and $L_{A^c,E}(x) = p(x|\omega) I_{A^c}(x) I_E(\omega)$.

E. Computing Reach Set Probabilities

Let us consider an agent as an SHA, H as in subsection III-A with the realization described by a Markov process M. Suppose that the hypothesis of subsection III-C are satisfied. In the remain of this section we propose some well-established statistical algorithms which might be employed in order to compute the reach set probabilities.

Solution 1. Let consider that we have given a credal set $\mathcal{K}$, which contains all probability densities of random variables x_t, $t \geq 0$. For a target set $A \in \mathcal{B}$, we define the measurable function

$$\varphi := \sup_{t \geq 0} I_A(x_t) \tag{16}$$

Consider μ an arbitrary element of $\mathcal{K}$ and $\underline{\mu}$, $\overline{\mu}$ given by (1). If $\underline{\mu}(E) > 0$, where E is an event in the sample probability space (a set of trajectories) the value of $\overline{\mu}(\varphi|E)$ can be computed by the generalized Bayes' rule (first proposed in [28]):

$$\overline{\mu}(\varphi|E) \text{ is the unique value of } \rho \text{ such that}$$
$$\overline{\mu}[(\varphi - \rho) I_E] = 0.$$

In this case, one can apply Lavine's algorithm, which is a bracketing scheme applied to the generalized Bayes' rule, whose objective is to compute upper expectations [20]. Define $\underline{\rho}_0 = \inf \varphi I_E$ and $\overline{\rho}_0 = \sup \varphi I_E$. Define $m(\rho) = \overline{\mu}[(\varphi - \rho) I_E]$; note that $m(\rho)$ must attain zero in the interval $[\underline{\rho}_0, \overline{\rho}_0]$. Now bracket this interval by repeating (for $i \geq 0$):
1. Stop if $|\overline{\rho}_i - \underline{\rho}_i| < \epsilon$ for some positive value ϵ; or
2. Choose ρ_i in $(\underline{\rho}_i, \overline{\rho}_i)$ and, if $m(\rho_i) > 0$, take $\underline{\rho}_{i+1} = \underline{\rho}_i$ and $\overline{\rho}_{i+1} = \overline{\rho}_i$; if $m(\rho_i) < 0$ take $\underline{\rho}_{i+1} = \underline{\rho}_i$ and $\overline{\rho}_{i+1} = \rho_i$. The value $m(\rho_i)$ can provide also information on when to stop the bracketing iteration [9].

The main inconvenient in applying Lavine's algorithm is that the function (16) is hard to compute. Another iteration scheme, also based on the generalized Bayes' rule, has been proposed by Walley [28] (Note 6.4.1). Both, Lavine and Walley algorithms have linear convergence.

Solution 2. We consider that the agent H has some beliefs w.r.t. the reach set probabilities, described by a $m-$close convex family $\mathcal{P}$ of probability measures on $(X, \mathcal{B})$ dominated by Cap_M. It should be clear that the credal set $\mathcal{P}$ depends on the random variables x_t, $t \geq 0$, which compose the process M. For example, $\mathcal{P}$ could contain the transition probabilities defined by (7).

Since the Choquet capacity associated to M (the realization of H) is the upper envelope of the family of probability measures dominated by Cap_M, the conditional reach set probability (14) is, in fact, obtained conditioning the upper probability $\overline{\mu}_M$ defined by (9).

Credal sets represented by supermodular (and dual by submodular) capacities have been studied in the literature (see [8] and the references therein). These credal sets have closed-form expressions for upper posterior probabilities. The algorithms, developed in [8], to compute the conditional upper (resp. lower) posterior probabilities are based on the Möbius Transform of an upper (resp. lower) probabilities. These algorithms seem to be the most suitable candidates which can be used to compute the reach set probabilities. In this paper, there is no room to give the necessary background for the understanding of these algorithms, and to present concretely how they can be used in the reachability analysis study. This will constitute the subject of a following paper.

There is a considerable debate about how an agent makes a decision using sets [1], [17]. We consider that an agent represented by an SHA should optimize the expected utility, an approach known in Robust Statistics under the name of Γ-minimax [17].

IV. Conclusions

The contribution of this paper is twofold:
1. We give a Bayesian framework for stochastic hybrid automata, which allow us to characterize the corresponding stochastic reachability problem in terms of upper probabilities or Choquet capacities.
2. Then the problem of computation of reach set probabilities finds new solutions using different Bayesian inference algorithms already studied in the literature.

In the first instance, this is a theoretical work, but practical implications of this approach are various and impressive. The benefits of this new representation of the stochastic reachability for ATC modelling are very important: Bayesian inference can be used to compute conflict probability and decision theory can be employed to derive algorithms for Conflict Resolution.

References

[1] Berger, J.O.: "Statistical Decision Theory and Bayesian Analysis". Springer-Verlag, 1985.
[2] Bensoussan, A., Menaldi, J.L.: Stochastic Hybrid Control. J. of Math. Analysis and Applications, **249** (2000), 261-288.
[3] Blumenthal, R.M., Getoor, R.K.: *Markov Processes and Potential Theory*. Academic Press, New York and London, 1968.
[4] Bujorianu, M.L., Lygeros, J.: *Reachability Questions in Piecewise Deterministic Markov Processes*. In O. Maler, A. Pnueli Eds., *Hybrid Systems: Computation and Control*, 6th International Workshop, HSCC03, LNCS 2623, pp.126-140, 2003.
[5] Bujorianu, M.L . *Extended Stochastic Hybrid Systems and their Reachability Problem*. In R. Alur, G. Pappas Eds., *Hybrid Systems: Computation and Control* 7th International Workshop, HSCC04 , pp.234-249, Springer LNCS 2993, 2004.
[6] Bujorianu, M.L., Lygeros, J.: *General Stochastic Hybrid Systems*. Proceedings of IEEE Mediterranean Conference on Control and Automation MED'04, 2004.
[7] Choquet, G.: *Theory of Capacities*. Annales de l'Institut Fourier, Grenoble, **5** (1953), 131-291.
[8] Chrisman, L.: Incremental Conditioning of Lower and Upper probabilities. Internat. J. Approx. Reason. **13**(1) (1995), 1-25.
[9] Cozman, F.G.: *Computing Posterior Upper Expectations*. Intern. J. Approx. Reasoning **24** (2000), 191-205.
[10] Davis, M.H.A.: "*Markov Models and Optimization*". Chapman & Hall, London, 1993.
[11] Dempster, D.: *Upper and Lower Probabilities Induced by a Multivalued Mapping*. Ann. Math. Statist. **38** (1967), 325-339.
[12] Fine, T.L.: *Lower Probability Models for Uncertainty and nondeterministic Processes*. J. Statist. Plann. Inference **20** (1988), 389-411.
[13] El Kaabouchi, A.: *Mesures Dominée par une Capacité Alternée D'ordre 2*. Proc. of the Amer. Math. Society, **121**(3) (1994), 823-832.
[14] Fitzsimmons, P.J.: *Markov Processes with Equal Capacities,* J. Theor. Prob. **12** (1999), 271-292.
[15] Getoor, R.K.: Excessive Measures. Birkhäuser, Boston, 1990.
[16] Huber, P.J., Strassen, V.: *Minimax Tests and the Neyman-Pearson Lemma for Capacities*. Ann. of Statist. **1** (1973), 251-263.
[17] Huber, P.J.: "Robust Statistics". Willey, New York, 1980.
[18] European Commission project HYBRIDGE, IST-2001-32460. Distributed Control and Stochastic Analysis of Hybrid Systems Supporting Safety Critical Real-Time System Design www.nlr.nl/public/hosted-sites/hybridge/
[19] Kyburg Jr., H.E.: *Bayesian and non-Bayesian Evidential Updating*. Artificial Inteligence **31** (1987), 271-293.
[20] Lavine, M.: Sensitivity in Bayesian Statistics the Prior and the Likelihood. J. Amer.statist. Assoc. **86** (414) (1991), 396-399.
[21] Levi, I.: "*The Enterprise of Knowledge*". MIT Press, Cambridge, Massachussetts (1980).
[22] Pola, G., Bujorianu, M.L., Lygeros, J., Di Benedetto, M. D.: *Stochastic Hybrid Models: An Overview with applications to Air Traffic Management*. ADHS, Analysis and Design of Hybrid System, Saint-Malo, 2003.
[23] Savage, L.J.: "*The Foundations of Statistics. Dover Publications*", Inc, New York, 1972.
[24] Shafer, G.: "*A Mathematical Theory of Evidence*". Princeton University Press, Princeton, New Jersey, 1976.
[25] Schmeidler, D.: *Subjective Probability and Expected Utility Without Additivity*. Econometrica **57** (1989), 571-587.
[26] Seidenfeld, T., Schervish, M., Kadane, J.B.: A Representation of Partially Ordered Preferences. Annal. Statist. **23** (6) (1995), 2168-2217.
[27] Suppes, P.: *The Measurement of Belief*. J. Roy. Statist. Soc. B 2 (1974), 160-191.
[28] Walley, P.: "*Statistical Reasoning with Imprecise Probabilities*". Chapman and Hall, London (1991).
[29] Wasserman, L.A., Kadane, J.B.: *Bayes's Theorem for Choquet Capacities*. The Annals of Statistics **18**(3) (1990), 1328-1339.