Stochastic Hybrid Models for Predicting the Behavior of Drivers Facing the Yellow-Light-Dilemma*

Daniel Hoehener[1], Paul A. Green[2] and Domitilla Del Vecchio[3]

***Abstract*— We address the problem of predicting whether a driver facing the yellow-light-dilemma will cross the intersection with the red light. Based on driving simulator data, we propose a stochastic hybrid system model for driver behavior. Using this model combined with Gaussian process estimation and Monte Carlo simulations, we obtain an upper bound for the probability of crossing with the red light. This upper bound has a prescribed confidence level and can be calculated quickly on-line in a recursive fashion as more data become available. Calculating also a lower bound we can show that the upper bound is on average less than 3% higher than the true probability. Moreover, tests on driving simulator data show that 99% of the actual red light violations, are predicted to cross on red with probability greater than 0.95 while less than 5% of the compliant trajectories are predicted to have an equally high probability of crossing. Determining the probability of crossing with the red light will be important for the development of warning systems that prevent red light violations.**

I. Introduction

In 2012 approximately 2.36 million people were injured in motor vehicle traffic crashes, about 30% of these injuries happened on or near signaled intersections [1]. Statistics show that driver distraction or inattention is the most prevalent contributing factor for all crashes at signaled intersections [2]. In [3], experiments have shown that using an on-board warning system, red light running could be reduced by 77%. Our objective is to design safety systems that are able to predict the probability of a red light violation. This ability will be used to issue warnings and if necessary, take control over the vehicle to prevent red light violations. In [4], [5], [6] safety systems were proposed for intersections without signals and by representing driver inputs as a disturbance. In this paper we present a combined experimental/theoretical study where suitably designed experiments in the driving simulator are used to create a stochastic model of driver behavior near signalized intersections. Considering the situation when the traffic light changes from green to yellow upon intersection approach (yellow-light-dilemma), we use this model to compute the probability of the driver being on the intersection while the traffic light is red.

Classification of driver behavior is an active area of research and several different approaches, mainly using machine learning techniques, have been proposed, see for instance [7], [8], [9] and the references therein. Most of these studies try to predict specific driving actions, such as turning left, going straight or stopping. In [7] the focus is on the prediction of traffic light violations and in [9] the authors suggest two methods to predict whether a driver is going to stop after observing a switch of the traffic light from green to yellow.

In this paper, instead, we seek to estimate the actual probability of reaching some given state (stochastic reachability problem). Moreover, we provide a complete model of driving behavior near intersections, which may be used for other purposes, including the design of safety-enforcing supervisors.

Stochastic reachability problems have been studied in the stochastic hybrid systems literature, see [10], [11], [12], [13]. Exact computation of reach probabilities remains a challenging problem but Monte Carlo methods have proven to be efficient, see [11].

By modeling driver behavior near intersections as a stochastic hybrid system, we make use of the existing stochastic reachability literature in order to formally define the probability of crossing on red. This probability is then computed by a combined Bayesian filter and Monte Carlo simulation approach. Tests on driving simulator data show the accuracy of the computed crossing probability. The method therefore provides a quantification of the danger instead of just a binary output (safe, dangerous).

In Section II, we state the problem and introduce the mathematical model. Then, in Section III, we present our solution algorithm and in Section IV we provide the experimental results.

II. System model and problem formulation

We first describe the intersection scenarios that we are considering and then introduce the mathematical model.

A. Application

When the traffic light changes from green to yellow, the driver has to decide whether he wants to brake or continue and try to make it through the intersection before the traffic light turns red. This is known as the yellow-light-dilemma. The question that we address is: What is the probability that the driver is going to be on the intersection while the traffic light is red? We shall call this the *crossing probability*.

The intersection scenarios that we consider start when the traffic light switches from green to yellow. The *vehicle state* $x = (p, v) \in \mathbb{R} \times \mathbb{R}_+$ is given by the vehicle position $p \in \mathbb{R}$, which represents the signed distance of the center of gravity

*This work was in part supported by NSF CPS Award number 1239182. The first author was supported by the SNSF.

[1]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA hoehener@mit.edu

[2]Transportation Research Institute, University of Michigan, Ann Arbor, MI 48109, USA pagreen@umich.edu

[3]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA ddv@mit.edu

of the subject vehicle to the intersection center and the subject vehicle's longitudinal speed $v \in \mathbb{R}_+$ (Figure 1). We

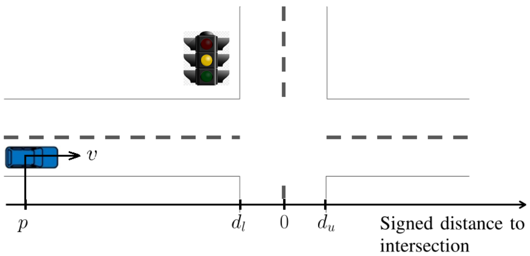

Fig. 1. Intersection with coordinate system.

assume that the vehicle state x is measured. In practice, this implies that the subject vehicle is equipped with differential GPS and a map of the area. In order to be able to estimate the crossing probability we make the additional assumption that the durations of the yellow and red light, denoted respectively by τ_y and τ_r, are known. The infrastructure should therefore be able to communicate these durations to the vehicle (V2I communication).

Figure 2 shows speed over position trajectories for different drivers once a yellow light is observed. We see clearly four different dynamical behaviors based on the driver's intended driving maneuver (action), namely coasting, braking, accelerating and waiting for the green light. That is, a different dynamical system belongs to each driver's action.

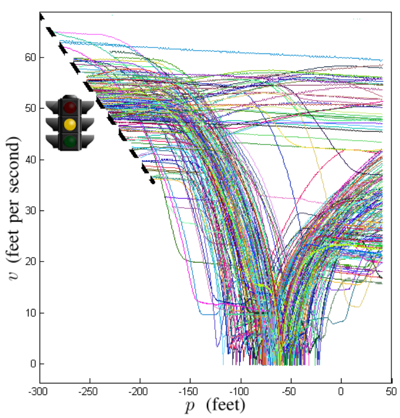

Fig. 2. Driver trajectories after observing a traffic light change.

Based on Figure 2, we seek a driver model with the following properties: 1) drivers have a finite set of basic actions that they can perform such as braking, coasting and waiting; 2) each basic action has its corresponding stochastic continuous dynamics; 3) the action intended by the driver is not directly observable. The framework of (general) stochastic hybrid systems [12] is adapted to build a driver model that satisfies these requirements. The next section introduces the mathematical definition of such a model.

B. Hybrid system model

We introduce here the stochastic hybrid system model and state some of its properties. A detailed treatment of the subject can be found for instance in [12].

a) State space: Hybrid systems have *continuous states*, that is, states that evolve according to a differential equation and discrete states, called *modes*, that evolve according to discrete transitions.

Let $n \in \mathbb{N}$ be given. The set of modes is denoted by Q and we define the set-valued map $\mathcal{X}\colon Q \rightsquigarrow \mathbb{R}^n$ which assigns to each mode $q \in Q$ an open set in $\mathbb{R}^n$. With this we define the *hybrid state space*,

$$\mathbf{X}(Q, \mathcal{X}) \coloneqq \bigcup_{q \in Q} \{q\} \times \mathcal{X}(q).$$

A *hybrid state* $(q, x) =: s \in \overline{\mathbf{X}(Q, \mathcal{X})}$ is a tuple formed by the mode $q \in Q$ and a continuous state $x \in \overline{\mathcal{X}(q)}$. Here $\overline{\mathcal{S}}$ denotes the closure of the set $\mathcal{S}$.

b) Formal definition and properties: Let $T > 0$ be a finite constant and $m \in \mathbb{N}$. Denote by $\{W_t\}_{t \in [0,T]}$ the m-dimensional standard Brownian motion, see for instance [14, Ch. 2]. A property used in this paper is that its increments are independent, normally distributed and have zero mean.

Definition 1: A (linear) *stochastic hybrid system* is a collection $H = (Q, \mathcal{X}, A, b, \sigma, R, Init, x_0)$ where Q and $\mathcal{X}\colon Q \rightsquigarrow \mathbb{R}^n$ are as above, $x_0 \in \mathbb{R}^n$ is the initial continuous state, $A\colon Q \to \mathbb{R}^{n \times n}$ and $\sigma\colon Q \to R^{n \times m}$ are matrix-valued maps, $b\colon Q \to \mathbb{R}^n$ is a vector-valued map, $R\colon Q \to [0, 1]$ is the transition measure and $Init\colon Q \to [0, 1]$ is the initial probability distribution on the modes.

The linear stochastic hybrid system defined here is a special case of a general stochastic hybrid system, see for instance [12, Def. 4.1], where dynamics are linear and only the modes can have jumps.

Definition 2: A stochastic process $\{\mathbf{s}(t)\}_{t \in [0,T]} = \{(\mathbf{q}(t), \mathbf{x}(t))\}_{t \in [0,T]}$ is an *execution* of a stochastic hybrid system H if there exist stopping times $T^0 = 0 \leq T^1 \leq \cdots \leq T^k \leq \cdots \leq T$ such that for each $k \in \{0, 1, \ldots\}$,

(i) $\mathbf{s}(0) = (q_0, x_0)$, where q_0 is a Q-valued random variable with probability distribution $Init$;
(ii) $\mathbf{q}(T^{k+1}) = q_{k+1}$, where q_{k+1} is a Q-valued random variable distributed according to R;
(iii) For all $t \in [T^k, T^{k+1})]$, where ")]" is closed when $T^{k+1} = T$ and open otherwise, $\mathbf{q}(t) \equiv q_k$;
(iv) For all $t \in]T^k, T^{k+1}]$, $\mathbf{x}(t)$ is the solution of

$$d\mathbf{x}(t) = (A(q_k)\mathbf{x}(t) + b(q_k))\, dt + \sigma(q_k) dW_t;$$

(v) $T^{k+1} = \inf \left\{ t \in]T^k, T] \;\middle|\; \mathbf{x}(t) \in \partial\mathcal{X}(q_k) \right\}$.

Throughout the paper $H = (Q, \mathcal{X}, A, b, \sigma, R, Init, x_0)$ represents a stochastic hybrid system which satisfies the standard assumptions [15, Assumption 1-3]. Other than regularity assumptions on the dynamics that are satisfied in the linear case, these assumptions demand that the executions have non-Zeno dynamics.

Definition 3: Let $C \in \mathbb{R}^{n \times 1}$ be given and $\{\mathbf{s}(t)\}_{t \in [0,T]}$ be an execution of H. An *output* of H corresponding to

$\{\mathbf{s}(t)\}_{t\in[0,T]}$ and C is the stochastic process defined by $\mathbf{y}(t) = C\mathbf{x}(t)$ for all $t \in [0, T]$.

Definition 4: A mode $q \in Q$ is *stationary* if $A(q) = b(q) = \sigma(q) = 0$.

Fact 1 ([15]): Every stochastic hybrid system H has an execution that is a strong Markov process (see [14, p. 81] for a definition).

Let $q \in Q$ be given and recall that the fundamental matrix $\Phi_q \colon [0, T] \to \mathbb{R}^{n\times n}$ of the ODE $\dot{x} = A(q)x + b(q)$ is

$$\Phi_q(t) \coloneqq e^{A(q)t} = \sum_{i=0}^{\infty} \frac{(A(q)t)^i}{i!}. \tag{1}$$

Fact 2: Let $q \in Q$ and $x \in \mathcal{X}(q)$. Then, the stochastic process $\{\mathbf{x}^q(t,x)\}_{t\in[0,T]}$ satisfying $\mathbf{x}^q(0,x) = x$ and

$$d\mathbf{x}(t,x) = (A(q)\mathbf{x}(t,x) + b(q))\, dt + \sigma(q) dW_t, \tag{2}$$

is given for all $t \in [0, T]$ by the stochastic integral

$$\mathbf{x}^q(t,x) = \Phi_q(t)\left(x + \int_0^t \Phi_q^{-1}(s) b(q) ds\right) + \Phi_q(t)\int_0^t \Phi_q^{-1}(s)\sigma(q) dW_s. \tag{3}$$

Moreover, $\{\mathbf{x}^q(t,x)\}_{t\in[0,T]}$ is a diffusion, i.e., time-homogenous and strongly Markovian with continuous sample paths. Finally, the stochastic process $\{E^q(t)\}_{t\in[0,T]}$ defined by

$$E^q(t) \coloneqq \Phi_q(t)\left(\int_0^t \Phi_q^{-1}(s)\sigma(q) dW_s\right), \tag{4}$$

is a Gaussian process, where for all $t, t' \in [0, T]$, $\mathbb{E}(E^q(t)) = 0$ and the covariance function $\Sigma^q(t,t')$ is given by

$$\Sigma^q(t,t') \coloneqq \int_0^{\min\{t,t'\}} \left(\Phi_q(t-s)\sigma(q)\sigma(q)^T\Phi_q(t'-s)^T\right) ds. \tag{5}$$

For a proof of these results see for instance [14] and [16]. The deterministic part of (3) is denoted by:

$$\varphi^q(t,x) \coloneqq \Phi_q(t)\left(x + \int_0^t \Phi_q^{-1}(s) b(q) ds\right). \tag{6}$$

C. Problem formulation

We assume the following data is available at all time.

Data: For $S \in [0, T]$ and $N^q \in \mathbb{N}$, $\forall q \in Q$, we have

- $I_t = [S, T]$ – time interval;
- $I_y = [y_{min}, y_{max}]$ – target set for the output;
- For each $q \in Q$, $\{e_1^q(t), \ldots, e_{N^q}^q(t)\}_{t\in[0,T]}$ is a set of observed sample paths of N^q independent stochastic processes distributed as $\{E^q(t)\}_{t\in[0,T]}$.

We impose the following assumptions on H.

Assumption 1: (i) There exists a stationary mode $\bar{q} \in Q$ and $\mathcal{X}(\bar{q}) = \mathbb{R}^n$; (ii) For all $q \in Q$ the transition measure is given by $R(q) = \mathbf{1}_{\{\bar{q}\}}(q)$, where $\mathbf{1}_{\mathcal{S}}(\cdot)$ is the indicator function of the set $\mathcal{S}$; (iii) There exists a set $\mathcal{T} \subset \mathbb{R}^n$ such that $\partial\mathcal{X}(q) = \mathcal{T}$ for all $q \in Q \setminus \bar{q}$.

The assumption implies that when the continuous state enters $\mathcal{T}$, then it must transit to the stationary mode $\bar{q}$. Moreover, there are no transitions out of the stationary mode.

Problem 1: Let $\alpha > 0$, $n \in \mathbb{N}$, $C \in \mathbb{R}^{n\times 1}$ and the above data be given. Moreover, let $\{\mathbf{s}(t)\}_{t\in[0,T]}$ be an execution of H and $\{\mathbf{y}(t)\}_{t\in[0,T]}$ the output corresponding to C. Finally, let $\mathbf{x}(t_0) = x_0, \ldots, \mathbf{x}(t_N) = x_N$ be measurements of the continuous state trajectory, where $0 = t_0 < t_1 < \cdots < t_N < T$. Find a $1-\alpha$ confidence upper bound for the probability

$$P_N \coloneqq \Pr(\mathbf{y}(I_t^N) \cap I_y \neq \emptyset \mid \mathbf{x}(t_0) = x_0, \ldots, \mathbf{x}(t_N) = x_N), \tag{7}$$

where $I_t^N \coloneqq I_t \cap [t_N, T]$.

Determining a $1-\alpha$ confidence upper bound implies that we seek an algorithm that will produce with probability $1-\alpha$ an upper bound for the true probability P_N. With an analogous approach we can find a $1-\alpha$ confidence lower bound to verify the tightness of the upper bound. The motivation for solving Problem 1 is to assess the risk of the output entering the set I_y. Taking an upper bound guarantees that this risk is not underestimated.

D. Illustration with application

Consider the application described in Section II-A. The time interval of interest is from the moment when the traffic light switches to yellow until it becomes green again. Hence $T \coloneqq \tau_y + \tau_r$ and time 0 is when the traffic light turns yellow. Modes represent the basic driver actions braking, coasting and waiting. Accelerating is not relevant to the model as it occurs after time T. Thus the set of modes is $Q = \{1, 2, 3\}$, where 1 stands for braking, 2 for coasting, and 3 for waiting. It is clear that waiting is a stationary mode. To determine the hybrid state space we have to consider that cars should not have negative speed and there should be a mode transition from braking to waiting when the car reaches zero speed. Recalling that the vehicle state $x = (p, v)$ is given by position p and speed v, we define $\mathcal{X} \colon Q \rightsquigarrow \mathbb{R}^2$:

$$\mathcal{X}(q) \coloneqq \begin{cases} \mathbb{R} \times]0, +\infty[& \text{if } q = \{1, 2\}, \\ \mathbb{R}^2 & \text{if } q = 3. \end{cases}$$

For the longitudinal dynamics we consider a second-order model. Thus for $q \in \{1, 2\}$, we have

$$A(q) = \begin{pmatrix} 0 & 1 \\ a_1^q & a_2^q \end{pmatrix},\ b(q) = \begin{pmatrix} 0 \\ b^q \end{pmatrix},\ \sigma(q) = \begin{pmatrix} 0 \\ \sigma_q \end{pmatrix}, \tag{8}$$

which also implies that we consider a one-dimensional standard Brownian motion. The Brownian motion models the uncertainty in the driver behaviors by introducing a random deviation from the nominal acceleration profile. We describe in Section IV-B how the parameters can be identified from data. The initial distribution of the modes $Init$ also has to be learned from data. The transition measure R is defined by Assumption 1.

Consider now the problem of estimating the crossing probability. The intersection is given by the interval $[d_l, d_u]$, see Figure 1. Thus we set $y_{min} \coloneqq d_l - d_f$ and $y_{max} =$

$d_u + d_r$, where d_f and d_r represent the distance from the vehicle's center of gravity to the vehicles front, respectively to its rear. Setting $S := \tau_y$, $I_t = [S, T]$ are the times when the traffic signal is red. With these definitions, the car is on the intersection if its position p is in the interval I_y. Hence we set $C = (1, 0)$.

III. Problem solution

Next we propose an algorithm to solve Problem 1. We start with a lemma that allows to decompose the problem into a mode estimation and a simpler reachability problem. Then we address these sub-problems.

A. Problem decomposition

Lemma 3.1: Let (t_i, x_i), $i \in \{0, \dots, N\}$, be as in Problem 1. Assume that $(t_i, Cx_i) \notin I_t \times I_y$ for all i. Then

$$P_N = \sum_{q \in Q} \Big(\Pr(\mathbf{q}(0) = q \mid \mathbf{x}(t_N) = x_N, \dots \mathbf{x}(t_0) = x_0) \cdot \\ \cdot \Pr(\mathbf{y}(I_t^N) \cap I_y \neq \emptyset \mid \mathbf{q}(t_N) = q, \mathbf{x}(t_N) = x_N) \Big), \quad (9)$$

if $x_N \notin \mathcal{T}$ and otherwise $P_N = \mathbf{1}_{I_y}(Cx_N)$.

To simplify the notation we abbreviate $\Pr(\mathbf{q}(0) = q \mid \mathbf{x}(t_N) = x_N, \dots \mathbf{x}(t_0) = x_0)$ by $\Pr(\mathbf{q}(0) = q \mid x_N, \dots, x_0)$.

Proof: By Assumption 1, if $x_N \in \mathcal{T}$ then $\mathbf{q}(t_N) = \bar{q}$ and $P_N = \mathbf{1}_{I_y}(Cx_N)$ follows from the stationarity of $\bar{q}$.

Consider now the case when $x_N \notin \mathcal{T}$. Then, by Assumption 1, we have that $t_N < \inf\{t \in]0, T] \mid \mathbf{x}(t) \in \mathcal{T}\}$, that is, no mode transition has occurred yet. In particular $\mathbf{q}(t_N) = \mathbf{q}(0)$. This leads to

$$P_N = \sum_{q \in Q} \Pr(\mathbf{q}(0) = q \mid x_N, \dots, x_0) \cdot \\ \cdot \Pr(\mathbf{y}(I_t^N) \cap I_y \neq \emptyset \mid \mathbf{q}(t_N) = q, x_N, \dots, x_0).$$

We then obtain (9) by using the Markov property of the execution $\{(\mathbf{q}(t), \mathbf{x}(t))\}_{t \in [0,T]}$, see Fact 1. ■

B. Mode estimation

Motivated by Lemma 3.1 we start by solving:

Problem 2: Let $x_0, \dots, x_N$ be as in Problem 1 and assume that $x_N \notin \mathcal{T}$. For all $q \in Q$ compute the probability

$$P_N^*(q) := \Pr(\mathbf{q}(0) = q \mid x_N, \dots, x_0). \quad (10)$$

The problem is solved by using Bayes' theorem. Denote by $f_{i,j}^q$ the joint density function of the random variables $\mathbf{x}(t_i), \mathbf{x}(t_{i+1}), \dots, \mathbf{x}(t_j)$ given $\mathbf{q}(0) = 0$, that is,

$$f_{i,j}^q(x) = f_{\mathbf{x}(t_i), \mathbf{x}(t_{i+1}), \dots, \mathbf{x}(t_j)}(x \mid \mathbf{q}(0) = q).$$

Analogously, for $i \in \{1, \dots, N\}$, define

$$f_i^q(x) = f_{\mathbf{x}(t_i)}(x \mid \mathbf{q}(0) = q).$$

Fix an arbitrary $q \in Q$. By Bayes' formula we have then

$$P_N^*(q) = \frac{f_{1,N}^q(x_1, \dots, x_N) Init(q)}{\sum_{\tilde{q} \in Q} f_{1,N}^{\tilde{q}}(x_1, \dots, x_N) Init(\tilde{q})}. \quad (11)$$

Recursive computation of (11) is achieved by the following update formulas which exploit the Markov property of executions, see Fact 1.

Update Formulas: For all $q \in Q$ we have that

$$P_1^*(q) = \frac{f_1^q(x_1) Init(q)}{\sum_{\tilde{q} \in Q} f_1^{\tilde{q}}(x_1) Init(\tilde{q})}. \quad (12)$$

Moreover for all $N > 1$,

$$P_N^*(q) = \frac{f_N^q(x_N \mid x_{N-1}) P_{N-1}^*(q)}{\sum_{\tilde{q} \in Q} f_N^{\tilde{q}}(x_N \mid x_{N-1}) P_{N-1}^*(\tilde{q})}. \quad (13)$$

C. Stochastic reachability of an interval

Let $\alpha \in]0, 1[$ be as in Problem 1 and define

$$\tilde{\alpha} := 1 - \sqrt[r-1]{1 - \alpha}, \quad (14)$$

where r is the number of modes. We solve the following problem:

Problem 3: For $q \in Q$ arbitrary, using the data provided in Section II-C, find a $1 - \tilde{\alpha}$ confidence upper bound for the probability

$$P_N(q) := \Pr(\mathbf{y}(I_t^N) \cap I_y \neq \emptyset \mid \mathbf{q}(t_N) = q, \mathbf{x}(t_N) = x_N).$$

Notice that for given $N \in \mathbb{N}$ and $q \in Q \setminus \bar{q}$, equations (12)-(13) provide an exact formula for the probability $P_N^*(q)$. Moreover, by solving Problem 3 we obtain a $1 - \tilde{\alpha}$ confidence upper bound for the probability $P_N(q)$, denoted by $u_N^q(\tilde{\alpha})$. The probability that $u_N^q(\tilde{\alpha}) \geq P_N(q)$ for all $q \in Q \setminus \bar{q}$ is then $(1 - \tilde{\alpha})^{r-1} = 1 - \alpha$. Since $P_N(\bar{q}) = \mathbf{1}_{I_y}(Cx_N)$, we have in this case the exact bound

$$u_N^{\bar{q}}(\tilde{\alpha}) = \mathbf{1}_{I_y}(Cx_N). \quad (15)$$

Consequently, by Lemma 3.1, a $1 - \alpha$ confidence upper bound for the probability P_N, is given by $\sum_{q \in Q} P_N^*(q) u_N^q(\tilde{\alpha})$.

To solve Problem 3 for $q \in Q \setminus \bar{q}$, we use N^q sample paths of the process $\{\mathbf{y}(t)\}_{t \in [0,T]}$ and then estimate for what fraction of the sample paths there exists a time in I_t^N for which the sample path takes values in I_y. By (3), we know that

$$\mathbf{x}(t) = \varphi^q(t - t_N, x_N) + E^q(t - t_N), \quad \forall t \in [t_N, T^1],$$

where $T^1 = \inf\{t \in [t_N, T] \mid \mathbf{x}(t) \in \mathcal{T}\}$. As $\varphi^q(t - t_N, x_N)$ is a deterministic function, we have that defining for all $i \in \{1, \dots, N^q\}$ the functions

$$x_i^q(t) := \varphi^q(t - t_N, x_N) + e_i^q(t), \quad \forall t \in [t_N, T],$$

$\{x_i^q(t)\}_{t \in [t_N, T]}^{i \in \{1, \dots, N^q\}}$ is a set of observed sample paths of independent processes distributed as $\{\varphi^q(t - t_N, x_N) + E^q(t - t_N)\}_{t \in [t_N, T]}$. This leads to the corresponding set of observed stopping times $T_i^1 := \inf\{t \in [t_N, T] \mid x_i^q(t) \in \mathcal{T}\}$. Finally, since by Assumption 1 $\mathbf{q}(t) = \bar{q}$, for all $t \geq T^1$, we infer that by defining

$$y_i^q(t) := \begin{cases} Cx_i^q(t) & \text{if } t < T_i^1 \\ Cx_i^q(T_i^1) & \text{otherwise,} \end{cases}$$

$\{y_i^q(t)\}_{t \in [t_N, T]}^{i \in \{1, \dots, N^q\}}$ is a set of observed sample paths of independent random processes $\{\mathbf{y}_i^q(t)\}_{t \in [t_N, T]}^{i \in \{1, \dots, N^q\}}$ identically distributed as $\{\mathbf{y}(t)\}_{t \in [t_N, T]}$ given that $\mathbf{s}(t_N) = (q, x_N)$.

Let us now associate with each of the random processes $\{\mathbf{y}_i^q(t)\}_{t\in[t_N,T]}^{i\in\{1,\ldots,N^q\}}$ a Bernoulli variable

$$Y_i^q := \begin{cases} 1 & \text{if } \mathbf{y}_i^q(I_t) \cap I_y \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

Since the processes $\{\mathbf{y}_i^q(t)\}$ are independent and identically distributed, the same is true for Y_i^q. Moreover, $P_N(q) = \Pr(Y_i^q = 1)$ for all i. It is well known that the sum $Z^q := \sum_{i=1}^{N^q} Y_i^q$ has the binomial distribution $B(N^q, P_N(q))$. Set for all $i \in \{1,\ldots,N^q\}$,

$$\gamma_i^q := \begin{cases} 1 & \text{if } y_i^q(I_t^N) \cap I_y \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

The values γ_i^q correspond to realizations of Y_i^q, hence $z^q := \sum_{i=1}^{N^q} \gamma_i^q$ is a realization of Z^q.

Providing a $1-\tilde{\alpha}$ confidence upper bound for the parameter P of the distribution $B(N^q, P)$ from a given observation is a standard statistical problem. For this study, we use the classical Clopper-Pearson confidence interval [17]. Hence a $1-\tilde{\alpha}$ confidence upper bound for $P_N(q)$ is given by,

$$u_N^q(\tilde{\alpha}) = Beta\left(1-\tilde{\alpha}; z^q+1, N^q-z^q\right), \quad (17)$$

where $Beta(\kappa; \upsilon, \nu)$ is the κ-quantile from a beta distribution with shape parameters υ and ν.

D. Solution algorithm

Using the results of the previous sections, we provide an algorithm to solve Problem 1.

Algorithm 1: 1. Let $N = 0$, $t_0 = 0$ and the initial state x_0 be observed. We define $\tilde{\alpha}$ as in (14) and initialize $P_0^*(q) := Init(q)$, for all $q \in Q$. By (9) we have that $\bar{P}_0 := \sum_{q\in Q} P_0^*(q) u_0^q(\tilde{\alpha})$, where $u_0^q(\tilde{\alpha})$ is defined by (15) and (17), is a $1-\alpha$ confidence upper bound for P_0.

2. For $N \to N+1$, let $t_{N+1} \in \,]t_N, T]$ be the current time and $\mathbf{x}(t_{N+1}) = x_{N+1}$ the new state observation. If $x_{N+1} \in \mathcal{T}$ or $t_{N+1} = T$, then $P_{N+1} = \mathbf{1}_{I_y}(Cx_{N+1})$ and the algorithm stops. Otherwise $P_{N+1}^*(q)$ is obtained from (12)-(13) and $\bar{P}_{N+1} := \sum_{q\in Q} P_{N+1}^*(q) u_{N+1}^q(\tilde{\alpha})$, where $u_{N+1}^q(\tilde{\alpha})$ is defined by (15) and (17). $\bar{P}_{N+1}$ is a $1-\alpha$ confidence upper bound for P_{N+1}. Repeat step 2.

IV. Application

A. Experimental setup

We use driving simulator data that was gathered at the University of Michigan Transportation Research Institute (UMTRI), see Figure 3. There were 24 subjects in this experiment. Twelve were under age 30 and 12 were older than 60.

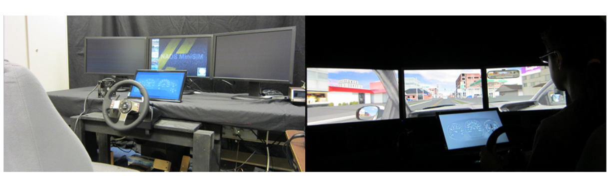

Fig. 3. Driving simulator at UMTRI where experiments were conducted

Within each age group there was an equal number of men and women. Each subject drove two test blocks, each block consisting of 70 intersections 200m apart. The subjects were instructed not to turn at any of the intersections (to make motion sickness less likely and simplify construction of the virtual world). In some intersections, the traffic light would remain green, in others it would turn to yellow and for some it would already be red upon approach. All intersections were crosses, with a single lane in each direction and in each intersection scenario there were up to four cars in addition to the subject vehicle.

In order to prevent excess speed, there was always a lead vehicle present, that is, a vehicle driving in front of the subject vehicle. The lead vehicle would however always cross the intersection when the traffic light changed to yellow, leaving the decision of whether to comply with the signal completely to the subject. Notice that the behavior of other traffic participants was not taken into account in the prediction (no vehicle-to-vehicle communication).

For our purpose, mainly intersections where the traffic light changed to yellow were of interest. In total, we considered $1,534$ such intersection approaches. The signal change to yellow would occur at three possible values for the time to intersection (TTI), which is the distance to the stop line divided by the current speed. Those values were respectively, 2.8s, 3.5s and 4.2s. The smaller TTI the faster subjects had to make a decision.

The data set provided by UMTRI includes position, speed and acceleration measurements for the subject vehicle as well as the traffic light information. Measurements were taken at a frequency of 60Hz.

B. System identification

This section is concerned with the identification of the model parameters A, b, σ and $Init$ from data. Since we used standard methods, we mainly provide references to the relevant literature.

As a first step, we divided the data into a training and a test set, both containing 767 intersection approaches. Only the training data was used for parameter identification. The separation into training and test data sets was done in a way that kept the ratios between male and female, old and young drivers, unchanged. Moreover, there was no overlap in subjects.

The training data was then further divided by identifying the trajectories belonging to the same mode. For all $q \in \{1,2\}$, the parameters a_1^q, a_2^q and b^q characterizing the maps given in (8) can then be identified in the same way as this was done in [18], i.e., by solving the least square optimization problem [18, (4)]. As a result we found the parameters $a_1^1 = -0.04$, $a_2^1 = -0.27$, $b^1 = -10.23$, $a_1^2 = -0.003$, $a_2^2 = 0.04$ and $b^2 = -2.12$.

The parameter σ_q defined in (8) is a so-called hyperparameter of a Gaussian process. Hyperparameters of Gaussian processes are classically estimated with a maximum likelihood method, see [19, Ch. 5]. The resulting parameters are $\sigma_1 = 2.54$ and $\sigma_2 = 0.66$.

Consider next the problem of identifying the initial distribution $Init$. Let $\{s(t)\}_{t\in[0,T]} = \{(q(t), p(t), v(t))\}_{t\in[0,T]}$ be a hybrid state trajectory from the training set. As described in Section IV-A, experiments were performed with the traffic light changing at three values for TTI, i.e. $p(0)/v(0) \in \{2.8, 3.5, 4.2\} =: I$. Define $\widetilde{Init}\colon Q \times I \to [0,1]$,

$$\widetilde{Init}(q,\tau) := \frac{\text{\# training trajectories s.t. } (q(0), \frac{p(0)}{v(0)}) = (q,\tau)}{\text{\# of training trajectories}}.$$

Then, using this and the law of large numbers, we have the estimator $\widehat{Init}(q; p_0, v_0) := \widetilde{Init}(q, p_0/v_0)$ for the initial distribution $Init$ of the hybrid system with initial state $x_0 = (p_0, v_0)$. In particular, we found that $\widetilde{Init}(1, 4.2) = 0.93$, $\widetilde{Init}(1, 3.5) = 0.81$, $\widetilde{Init}(1, 2.8) = 0.47$, $\widetilde{Init}(2,\tau) = 1 - \widetilde{Init}(1,\tau)$ for all $\tau \in I$ and finally $\widetilde{Init}(3,\tau) = 0$ for all $\tau \in I$. These values show that less drivers will brake when time to intersection decreases.

When the traffic light changes the driver's reaction to the light switch is not reflected immediately in position and speed measurements. Therefore we start Algorithm 1 only 2s after the light change. The 2s value corresponds to the 90% quantile of the cumulative human response time distribution, see [20]. Response time was defined as the time from the moment the risk is presented to the driver until the driver input starts, see the SAE J2944 standard.

C. Experimental results

In this section, we provide results obtained from Algorithm 1. The algorithm was implemented in MATLAB and run on a 2.6GHz dual-core computer. To compute the mode update, notice that in our application for $N \le 20$, $\Pr(\exists t \in [0, t_N], \mathbf{x}(t) \in \mathcal{T} \mid \mathbf{q}(0) = q) \approx 0$ for all $q \in Q$. This is because the 2s corresponding to $N = 20$ are too short to decelerate the car to zero speed with a normal braking maneuver. Hence $f_i^q(x) \approx f_{\mathbf{x}^q(t_i)}(x)$, for all $i \in \{1, \dots, N\}$, which is a Gaussian density. Thus computing the mode update using (12)-(13) only requires to evaluate Gaussians and takes less than 0.3ms. Computing $u_N^q(\tilde{\alpha})$, takes less than 5ms. A full iteration of Algorithm 1 is performed in less than 10ms. All results were obtained by running the algorithm on the 767 intersection approaches of the test data set and with parameter $\alpha = 0.05$. 478 of these approaches comply with the traffic light, the remaining 289 are violating trajectories.

The purpose of Algorithm 1 is to assess the risk that a car will cross on red by providing an $1-\alpha$ confidence upper bound. Using an analogous procedure as to compute the upper bound $\bar{P}_N$, we can compute a corresponding $1-\alpha$ confidence lower bound, that we denote by $\tilde{P}_N$. It follows that $\left|\bar{P}_N - P_N\right| \le \bar{P}_N - \tilde{P}_N$ with probability $1-2\alpha$. Table I shows the average difference $\bar{P}_N - \tilde{P}_N$ as a function of the number of observations N. Measurements were taken at a frequency of 10Hz and the average is taken over all 767 trajectories from the test set.

Table I shows that prediction accuracy increases slowly with the number of measurements and that independent of the number of measurements N, $\bar{P}_N - P_N$ is on average less than 0.023. The standard deviation is always less than $4 * 10^{-4}$. This bound is theoretical in the sense that it is based on the theoretical result that for all q, $P_N^*(q)$ is the actual probability of mode q. It is, however, confirmed by our experiments. We ran Algorithm 1 on each test trajectory and made predictions at a frequency of 10Hz, which led to a total of 14,623 predictions, of which at most 20 were taken from the same trajectory. In 98% of the 5,301 cases when $\bar{P}_N$ was larger than 0.95, the vehicle would actually cross on red. Similarly, in less than 1% of the 7,979 cases when $\bar{P}_N$ was lower than 0.05, the vehicle would cross on red.

TABLE I

TIGHTNESS OF UPPER BOUND $\bar{P}_N$

	Number of observations N			
	1	5	10	15
Avg. of $\bar{P}_N - \tilde{P}_N$	0.023	0.021	0.021	0.02

A crucial question from an application point of view is how many observations are needed to predict a traffic light violation with high probability, assuming there will be one. To be more precise, call a prediction *decisive* when the crossing probability is above 0.95 and then define the *detection rate* at a given time as the percentage of traffic light violating trajectories that have gotten a decisive prediction at that time. Table II compares the detection rates for the algorithm if we take measurements and update the probability at 5, 10 and 30Hz respectively. Data was obtained by running the algorithm on all 289 traffic light violating intersection approaches.

TABLE II

DETECTION RATES OF RED CROSSING TRAJECTORIES

	Elapsed time in seconds				
	0.033	0.067	0.1	0.2	0.4
Detection rate 30Hz	51	80	92	99	99
Detection rate 10Hz	–	–	84	96	99
Detection rate 5Hz	–	–	–	92	98

As the results in Table II show, the traffic light violations are detected by the algorithm in most cases in less than 0.2s.

In the recent paper [7] the problem of detecting traffic light violations was studied. In order to give traffic participants time to react, it was required that warnings are given (if necessary) before TTI becomes smaller than a lower bound $TTI_{min} > 0$, see Section IV-A. We use the same values as in [7], i.e., $TTI_{min} \in \{1s, 1.6s, 2s\}$, corresponding to the human response time distribution percentiles 45%, 80% and 90% respectively, see [20]. Table III shows the result for our red light crossing prediction. We say that crossing is detected at time TTI_{min} whenever it has a decisive prediction. In addition to detection rate, the table shows the percentage of compliant trajectories that were classified as crossing, these are called false positives. Finally, the last row shows the percentage of violating trajectories within the trajectories that were classified as dangerous, this is called the *percentage of justified warnings*. As in [7], position measurements were taken at 10Hz during a maximum of 2s or until the bound on TTI was reached. To allow the drivers to respond to the yellow light before TTI would become smaller than TTI_{min}, we used only the 204 trajectories where the traffic light changed when TTI was 4.2s. 27 of these trajectories were traffic light violations.

TABLE III
DETECTION AND FALSE POSITIVE RATES AT CRITICAL TTI VALUES

	TTI_{min}		
	1s	1.6s	2s
% Detected actual crossing	96	96	81
% Falsely detected crossing	0	2	4
% Justified warnings	100	87	76

We see in Table III that the detection rate is high, even with the largest TTI_{min}, while unjustified warnings remain on an acceptable level (24%).

A major difference in our scenario compared to [7] is that we start the algorithm when the traffic light turns yellow. Consequently, the time window until TTI_{min} varies from case to case, while in [7] the number of observations available to perform the classification was fixed. Moreover, we consider a scenario with a traffic light change while in [7] there is always a red light. The detection rate in this study is at least 10% higher in all cases, the false positive rate is always below the 5% of [7] and even in the worst case we have 76% justified warnings, while even in the best case in [7] it is only 63%[1].

In [9] the authors proposed two algorithms to predict whether a car would cross after observing a yellow light. The methods were compared with those of [7]. For $TTI_{min} = 1$s the detection rate in [9] is 100%, however for the other two cases our detection rate is at least 5% higher and we have lower false positive rates in all cases. Finally, notice that the methods in [7] and [9] use also acceleration measurements while Algorithm 1 does not.

V. CONCLUSIONS

We studied the dilemma a driver is facing when the traffic light changes to yellow. Our objective was to determine an upper bound on the crossing probability, having a prescribed confidence level. The algorithm presented here is based on a stochastic hybrid system model with hidden modes and uses Gaussian process theory to estimate the mode online using measurements of the continuous state only. For testing we used 767 intersection approaches recorded during experiments in a driving simulator. We find that the percentage of actual crossing trajectories within the set of trajectories that were predicted to cross with a probability smaller than 0.05 was 1%. Similarly, the percentage of actual crossing trajectories within the set of trajectories that were predicted to cross with probability larger than 0.95 was 98%. Moreover, the percentage of crossing trajectories that were predicted to cross with probability larger than 0.95 within the set of all actual crossing trajectories is 99%. These results show the accuracy of the predictions and that in most cases crossing trajectories can be identified.

An important direction for future research is the use of the constructed model to design warning/override systems to prevent red light violations and warn other traffic participants of dangerous situations.

[1] This value is inferred from [7, Table IV] and the fact that there were 8,000 compliant and 800 violating trajectories [7, Section V].

REFERENCES

[1] U.S. Department of Transportation National Highway Safety Administration. (2013, November) 2012 motor vehicle crashes: Overview. [Online]. Available: http://www-nrd.nhtsa.dot.gov/Pubs/811856.pdf

[2] P. Green, C. Demeniuk, and S. Jih, "In-vehicle traffic signal violation warnings: A review of the human factors literature," University of Michigan Transportation Research Institute, Tech. Rep., 2008.

[3] P. Green, J. Schweitzer, M. Alter, and C. Demeniuk, "Traffic signal violation warnings: Driver interface development and an initial driving simulator evaluation," University of Michigan Transportation Research Institute, Tech. Rep., 2008.

[4] M. Hafner and D. Del Vecchio, "Computational tools for safety control of a class of piecewise continuous systems with imperfect information on a partial order," *SIAM J. Control Optim.*, vol. 49, pp. 2463–2493, 2011.

[5] R. Verma and D. Del Vecchio, "Safety control in hidden mode hybrid systems," *IEEE Trans. Automatic Control*, vol. 57, pp. 62–77, 2012.

[6] ——, "Semiautonomous multivehicle safety," *IEEE Robotics and Automation Magazine*, vol. 18, pp. 44–54, 2011.

[7] G. S. Aoude, V. R. Desaraju, L. H. Stephens, and J. P. How, "Driver behavior classification at intersections and validation on large naturalistic data set," *IEEE Trans. Intelligent Transportation Systems*, vol. 13, pp. 724–736, 2012.

[8] V. Gadepally, A. Krishnamurthy, and U. Ozguner, "A framework for estimating driver decisions near intersections," *IEEE Trans. Intelligent Transportation Systems*, pp. 637–646, 2014.

[9] R. Mabuchi and K. Yamada, "Prediction of driver's stop or go at yellow traffic signal from vehicle behavior," in *IEEE Intelligent Vehicles Symposium (IV)*, 2013.

[10] A. Abate, S. Amin, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability for discrete time stochastic hybrid systems," in *Proc. IEEE Conference on Decision and Control*, 2006, pp. 258–263.

[11] H. A. P. Blom, G. J. B. Bakker, and J. Krystul, "Probabilistic reachability analysis for lager scale stochastic hybrid systems," in *Proc. IEEE Conference on Decision and Control*, 2007, pp. 3182–3189.

[12] M. L. Bujorianu, *Reachability Analysis of Hybrid Systems*. Springer-Verlag London, 2012.

[13] M. Prandini and J. Hu, "Application of reachability analysis for stochastic hybrid systems to aircraft conflict prediction," in *Proc. IEEE Conference on Decision and Control*, 2008, pp. 4036–4041.

[14] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*. Springer-Verlag, 1991.

[15] M. L. Bujorianu and J. Lygeros, "Toward a general theory of stochastic hybrid systems," in *Stochastic Hybrid Systems*, ser. Lecture Notes in Control and Information Science, H. A. Blom and J. Lygeros, Eds. Springer Berlin Heidelberg, 2006, vol. 337, pp. 3–30.

[16] B. Øksendal, *Stochastic differential equations*. Springer-Verlag, 1985.

[17] C. J. Clopper and E. S. Pearson, "The use of confidence for fiducial limits illustrated in the case of the binomial," *Biometrika*, vol. 26, pp. 404–413, 1934.

[18] D. Del Vecchio, R. M. Murray, and P. Perona, "Primitives for human motion: A dynamical approach," in *Proc. IFAC World Congress, Barcelona*, 2002.

[19] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[20] S. McLaughlin, J. Hankey, and T. Dingus, "A method for evaluating collision avoidance systems using naturalistic driving data," *Accident Anal. Prev.*, vol. 40, pp. 8–16, 2008.