# Safety Control of a Class of Stochastic Order Preserving Systems with Application to Collision Avoidance near Stop Signs

Mojtaba Forghani[1], John M. McNew[2], Daniel Hoehener[1] and Domitilla Del Vecchio[1]

***Abstract*— In this paper, we consider the problem of keeping the state of a system outside of an undesired set of states with probability at least *P*. We focus on a class of order preserving systems with a constant input disturbance that is extracted from a known probability distribution. Leveraging the structure of the system, we construct an explicit supervisor that guarantees the system state to be kept outside the undesired set with at least probability *P*. We apply this supervisor to a collision avoidance problem, where a semi-autonomous vehicle is engaged in preventing a rear-end collision with a preceding human-driven vehicle, while stopping at a stop sign. We apply the designed supervisor in simulations in which the preceding vehicle trajectories are taken from a test data set. Using this data, we demonstrate experimentally that the probability of preventing a rear-end collision while stopping at the stop sign is at least *P*, as expected from theory. The simulation results further show that this probability is very close to *P*, indicating that the supervisor is not conservative.**

## I. INTRODUCTION

The problem of designing control strategies that guarantee the safety of a system, that is, avoidance of a dangerous set of states, has been studied for many years in the context of deterministic systems, chiefly by [1]–[3]. This problem has been solved by deriving the Hamilton-Jacobi-Bellman equation whose solutions describe the boundary of the maximal safe controlled invariant set. While in general computing this set is computationally difficult, a number of ground transportation systems can be modeled by a special class of systems, called order preserving systems, that allow computationally efficient solutions [4]–[8].

When the system model is stochastic, the problem of designing safety-enforcing controllers has been addressed only more recently. In particular, the maximum achievable safety probability for a given initial state for stochastic nonlinear and hybrid systems has been investigated in [9]. In [10] and [11], the corresponding control policy that guarantees this maximal safety probability is also provided. Safety for a given probability *P*, which is our primary goal in this work, for a particular class of systems has been addressed in [20]. This problem is of practical relevance in a number of application scenarios, including the design of on-board driver-assist systems that warn/override the driver to guarantee a prescribed safety level. The application of our algorithms to collision avoidance scenarios near stop signs is the second goal of our work.

The problem of preventing or mitigating collisions near intersections (signaled or not) is a major focus of research due to the large number of collisions and fatalities that still occur today world-wide [12]. For example, in the United States, over the last several years an average of 21% of the fatalities and roughly 50% of the serious injuries have been attributed to intersections [13]. In order to design driver-assist systems that apply a warning or an override at the right time when the surrounding vehicles do not communicate, it is important to have a model of the behavior of these vehicles. Previously, deterministic models were considered for controller design, wherein the vehicle behavior was modeled through a set of modes with bounded disturbances capturing the variability among and within drivers in each mode (see [4] and [5]). Since driver's behavior and variability among drivers is better captured by probabilistic disturbances, we consider here a model where the disturbance has a probability distribution, which can be learned from data. This allows to design warnings and overrides that are less conservative and can guarantee a given probability of safety.

In [14] and [15], HMMs (Hidden Markov Model) were employed as a stochastic model for driver behavior for estimation/prediction purposes. While these models provide the desired results for estimation/prediction tasks, given their complexity, they are less suited for real-time control purposes. In this paper, we therefore consider a simpler model in which the continuous dynamics are order preserving and the disturbance inputs are constant parameters distributed according to a Gaussian probability distribution. With these assumptions, we provide a control map that can be efficiently computed on-line for guaranteeing a given probability of safety *P*. We apply our algorithms to a collision avoidance scenario wherein one vehicle needs to stop at a stop sign while preventing a collision with a preceding human-driven vehicle. The model of the preceding vehicle is learned from data gathered from vehicles driving in Ann Arbor (MI). A different data set for the preceding vehicle trajectories was used in simulations to emulate the preceding vehicle. In these simulations, the following vehicle was supervised through our control algorithm that provided overrides to ensure a probability of safety *P*. Simulation results show that the prescribed probability of safety *P* (and not more than *P*) was indeed ensured validating the algorithms on experimental data and demonstrating the non-conservatism of the approach.

This paper is organized as follows. In Section II, we pro-

This work was supported by NSF under the award #1161893

[1] Mojtaba Forghani, Daniel Hoehener and Domitilla Del Vecchio are with the Department of Mechanical Engineering, MIT, 77 Massachusetts Avenue, Cambridge, MA. mojtaba@mit.edu, hoehener@mit.edu, ddv@mit.edu

[2] John M. McNew is with Toyota Technical Center, 1555 Woodridge Avenue, Ann Arbor, MI. john.mcnew@tema.toyota.com

vide details of the stochastic model, introduce the collision avoidance application, and formulate the control problems. In Section III, we solve the control problems and in Section IV we provide the details of the implementation.

## II. STOCHASTIC MODEL

### A. System Model

We start with some basic definitions.

**Definition 1:** For all $w,z\in\mathbb{R}^n$ we have that $w\leq z$ $(w<z)$, if and only if $w_i\leq z_i$ $(w_i<z_i)$ for all $i\in\{1,2,...,n\}$, in which $w_i$ denotes the $i$th component of $w$. We denote the piecewise continuous signal on $\mathscr{U}$ by $S(\mathscr{U}):\mathbb{R}_+\rightarrow\mathscr{U}$. For $\mathscr{U}\subset\mathbb{R}^m$ we define the partial order (strict partial order) by component-wise ordering for all times, that is, for all $\mathbf{w},\mathbf{z}\in S(\mathscr{U})$ we have that $\mathbf{w}\leq\mathbf{z}$ $(\mathbf{w}<\mathbf{z})$ provided $\mathbf{w}(t)\leq\mathbf{z}(t)$ $(\mathbf{w}(t)<\mathbf{z}(t))$ for all $t\in\mathbb{R}_+$. The map $f:P\rightarrow Q$ is order preserving (strict order preserving) provided if for $x,y\in P$ we have $x\leq y$ $(x<y)$, then $f(x)\leq f(y)$ $(f(x)<f(y))$.

**Definition 2:** A continuous system is a collection $\Sigma=(X,U,\Delta,O,f,h)$, with state $x\in X\subset\mathbb{R}^n$, control input $u\in U\subset\mathbb{R}^m$, disturbance input $d\in\Delta\subset\mathbb{R}^q$, output $y\in O\subset X$, vector field in the form of $f:X\times U\times\Delta\rightarrow X$, and output map $h:X\rightarrow O$.

**Definition 3:** For $\Sigma^1=(X^1,U^1,\Delta^1,O^1,f^1,h^1)$ and $\Sigma^2=(X^2,U^2,\Delta^2,O^2,f^2,h^2)$, we define the parallel composition $\Sigma=\Sigma^1||\Sigma^2:=(X,U,\Delta,O,f,h)$, in which $X=X^1\times X^2$, $U:=U^1\times U^2$, $\Delta:=\Delta^1\times\Delta^2$, $O:=O^1\times O^2$, $f:=(f^1,f^2)$ and $h:=(h^1,h^2)$.

We denote the flow of a system $\Sigma$ at time $t\in\mathbb{R}_+$ by $\phi(t,x,\mathbf{u},\mathbf{d})$, with initial condition $x\in X$, control input signal $\mathbf{u}\in S(U)$, and disturbance input signal $\mathbf{d}\in S(\Delta)$. We also denote the $i$th component of the flow by $\phi_i(t,x,\mathbf{u},\mathbf{d})$.

**Definition 4:** A continuous system $\Sigma=(X,U,\Delta,O,f,h)$ is called input/output order preserving (strict input/output order preserving) with respect to the control input signal, if the map $h(\phi(t,x,\cdot,\mathbf{d})):S(U)\rightarrow O$, for any fixed $t$, $x$ and $\mathbf{d}$, is order preserving (strict order preserving).

**Definition 5:** A continuous system $\Sigma=(X,U,\Delta,O,f,h)$ is called input/output order preserving (strict input/output order preserving) with respect to the disturbance input signal, if the map $h(\phi(t,x,\mathbf{u},\cdot)):S(\Delta)\rightarrow O$, for any fixed $t$, $x$ and $\mathbf{u}$, is order preserving (strict order preserving).

In this paper, we consider system $\Sigma^*=\Sigma^1||\Sigma^2$, which is the parallel composition of $\Sigma^1=(X^1,U,\emptyset,O^1,f^1,h^1)$ and $\Sigma^2=(X^2,\emptyset,\Delta,O^2,f^2,h^2)$, where $x^1\in X^1\subset\mathbb{R}^{n^1}$, $x^2\in X^2\subset\mathbb{R}^{n^2}$, $u\in U=[u_m,u_M]\subset\mathbb{R}^m$, with $u_m\in\mathbb{R}^m$ and $u_M\in\mathbb{R}^m$ the minimal and the maximal control inputs for $\Sigma^1$, respectively, $d\in\Delta=\mathbb{R}$, $y^1\in O^1$, $y^2\in O^2$, $h^1:X^1\rightarrow O^1$, $h^2:X^2\rightarrow O^2$, $f^1:X^1\times U\rightarrow X^1$ and $f^2:X^2\times\Delta\rightarrow X^2$. Since $\Delta^1=\emptyset$ and $U^2=\emptyset$, we represent the flows of systems $\Sigma^1$ and $\Sigma^2$ by $\phi^1(t,x^1,\mathbf{u})$ and $\phi^2(t,x^2,\mathbf{d})$, respectively. The following assumptions are made on system $\Sigma^*$.

**Assumption 1:** System $\Sigma^1$ is input/output order preserving with respect to the control input and its flow $\phi^1$ is continuous in all arguments.

**Assumption 2:** System $\Sigma^2$ is strict input/output order preserving with respect to the disturbance input and its flow $\phi^2$ is continuous in all arguments.

**Assumption 3:** The disturbance input is a constant with Gaussian distribution, that is, $\mathbf{d}(t):=d\sim\mathcal{N}(\mu,\sigma^2)$, for all $t\in\mathbb{R}_+$.

### B. Application Scenario

We consider the scenario of two consecutive vehicles approaching a stop sign. We assume that the following vehicle (FV) is equipped with the collision avoidance system, while the preceding vehicle (PV) is fully human driven. We consider two types of "collisions": type (1), the rear-end collision between the two vehicles; type (2), crossing the stop sign with a high velocity. We denote longitudinal position and velocity of PV by $x_p$ and $v_p$, respectively. Similarly, $x_f$ and $v_f$ are position and velocity of FV, respectively. The longitudinal position of the stop sign is $St$ and the maximum allowable velocity of FV at the stop sign is $v_T$. The minimum allowable distance between the two vehicles is $\delta>0$. The scenario is depicted in Figure 1.

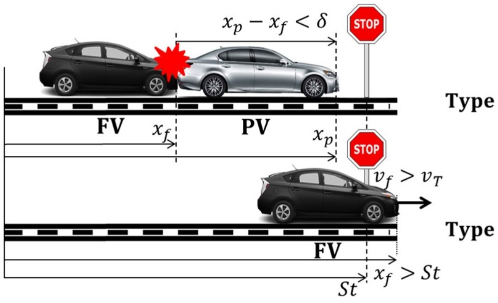


Fig. 1: Collision scenarios.

The system model for the application scenario is given by $\Sigma_{app}:=\Sigma^1||\Sigma^2$, where $\Sigma^1$ and $\Sigma^2$ are FV and PV, respectively. Hence, $x^1=(x_f,v_f)^T$, $x^2=(x_p,v_p)^T$, $y^1=x_f$, $y^2=x_p$, $h^1(x^1)=x_f$ and $h^2(x^2)=x_p$. The deceleration due to the rolling resistance and the slope of the road of FV are $a_r$ and $a_s$, respectively. We let $D$ denote the drag coefficient. We also assume that the speed of both vehicles is non-negative. The control input is $u\in U\subset\mathbb{R}$, and the disturbance input is $d\in\mathbb{R}$. We define functions $f^1(x^1,\mathrm{u}^1)$ and $f^2(x^2,\mathrm{u}^2)$, where $\mathrm{u}^1=u$ and $\mathrm{u}^2=d$, as follows.

$$\text{for } i\in\{1,2\},\ f^i(x^i,\mathrm{u}^i)=\begin{cases}\bar{f}^i(x^i,\mathrm{u}^i) & \text{if } \mathrm{v}^i>0\\ 0 & \text{if } \mathrm{v}^i\leq 0\end{cases}, \quad (1)$$

where $\mathrm{v}^1=v_f$ and $\mathrm{v}^2=v_p$. Also, $\bar{f}^1(x^1,u)$ and $\bar{f}^2(x^2,d)$ are

$$\bar{f}^1(x^1,u)=\left(v_f,\ u-Dv_f^2-a_r-a_s\right)^T, \quad (2a)$$

$$\bar{f}^2(x^2,d)=(v_p,\ ax_p+bv_p+d)^T. \quad (2b)$$

The term $ax_p+bv_p+d$ is the acceleration of PV. More details on this model are provided in Section IV. We assume that $d\sim\mathcal{N}(\mu,\sigma^2)$, which is consistent with Assumption 3.

Based on Assumption 1, the flow $h^1(\phi^1(t,x^1,\mathbf{u}))=\phi^1_1(t,x^1,\mathbf{u})=x_f(t)$ must be order preserving with respect to

$\mathbf{u}$. In the following proposition, we prove that both $x_f(t)$ and $\phi_2^1(t,x^1,\mathbf{u}) = v_f(t)$ are order preserving with respect to $\mathbf{u}$.

**Proposition 1:** The flows $\phi_1^1(t,x^1,\mathbf{u})$ and $\phi_2^1(t,x^1,\mathbf{u})$ of $\Sigma_{app}$ are order preserving with respect to the control input.

*Proof:* (Sketch) We consider two different control input signals $\mathbf{u}_1$ and $\mathbf{u}_2$ such that $\mathbf{u}_1 > \mathbf{u}_2$. Then using equation (2a) and continuity of the flow with respect to time we can prove that for the velocity of FV at time $t$ corresponding to $\mathbf{u}_1$ and $\mathbf{u}_2$ starting from the same initial condition, denoted by $v_{f,1}(t)$ and $v_{f,2}(t)$, respectively, we have $v_{f,1}(t) - v_{f,2}(t) \geq 0$. Since $x_{f,1}(t) - x_{f,2}(t) = \int_0^t v_{f,1}(s) - v_{f,2}(s)ds$, then also $x_{f,1}(t) - x_{f,2}(t) \geq 0$ (see [19] for more details). ■

In Proposition 2, we prove that Assumption 2 is also valid for our application scenario, that is, $x_p(t) = h^2(\phi^2(t,x^2,\mathbf{d})) = \phi_1^2(t,x^2,\mathbf{d})$ is strictly order preserving with respect to $\mathbf{d}$.

**Proposition 2:** The flow $\phi_1^2(t,x^2,\mathbf{d})$ of $\Sigma_{app}$ is strictly order preserving with respect to the disturbance input.

*Proof:* (Sketch) From the definition of $\bar{f}^2(x^2,d)$ in (2b) we have that the velocity of PV, for $v_p(t) > 0$, satisfies the differential equation $\ddot{v}_p - b\dot{v}_p - av_p = 0$. We consider two disturbance signals $\mathbf{d}^1$ and $\mathbf{d}^2$ such that $\mathbf{d}^1 > \mathbf{d}^2$, and then by solving the differential equation we can prove that for any $t \in \mathbb{R}_+$, for the velocity at time $t$ corresponding to $\mathbf{d}^1$ and $\mathbf{d}^2$, denoted by $v_{p,1}(t)$ and $v_{p,2}(t)$, respectively, we have $v_p^1(t) - v_p^2(t) > 0$ (see [19] for more details). Since $x_{p,1}(t) - x_{p,2}(t) = \int_0^t v_{p,1}(s) - v_{p,2}(s)ds > 0$, then $x_p$ is strictly order preserving with respect to the disturbance input $\mathbf{d}$. ■

### C. Problem Formulation

We use $Pr(\cdot)$ to denote the probability. We use $\dim S$ to denote the dimension of a vector space $S$. The $i$th row and $j$th column of a matrix $A$ is denoted by $A_{ij}$. A static feedback map is represented by $\pi : X \to U$, where $\mathbf{u}(t) = \pi(x(t))$. For a set $S \subset X$, we define $S^c := \{x \in X \mid x \notin S\}$.

**Assumption 4:** Bad set is in the form $B = B_1 \cup B_2$, where

$$B_1 = \bigcup_{j=1}^{N} \left\{x \in X \;\middle|\; G^j(x^1) > g^j \right\} \text{ and}$$

$$B_2 = \left\{x \in X \;\middle|\; Z^1 h^1(x^1) - Z^2 h^2(x^2) > H \right\},$$

where $Z^1$ and $Z^2$ are non-negative $r \times \dim(O^1)$ and $r \times \dim(O^2)$ matrices, respectively, $H$ is a $r$-dimensional vector, $G^j : X^1 \to \mathbb{R}^{p^j}$ and $g^j$ is a $p^j$-dimensional vector.

**Assumption 5:** For fixed $t \in \mathbb{R}_+$, $x^1 \in X^1$, and $j \in \{1,...,N\}$, $G^j(\phi^1(t,x^1,\cdot)) : S(U) \to \mathbb{R}^{p^j}$ is order preserving.

The two following problems concerned with the probabilistic safety of system $\Sigma^*$ must be solved.

**Problem 1:** For system $\Sigma^*$, with Assumptions 1-5 and $P \in (0,1)$, find the open loop maximal safe set given by

$$\mathscr{W} := \{x \in X \mid \exists\; \mathbf{u} \in S(U) \; s.t. \; Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B, \\ \forall t \in \mathbb{R}_+ \text{ and } \mathbf{d}(t) \sim \mathscr{N}(\mu,\sigma^2)) \geq P\}.$$

**Problem 2:** For system $\Sigma^*$, with Assumptions 1-5 and $P \in (0,1)$, find a control map $\pi : X \to U$ such that for all $x \in \mathscr{W}$ we have $Pr(\forall t \in \mathbb{R}_+ \text{ and } \mathbf{d}(t) \sim \mathscr{N}(\mu,\sigma^2), \phi(t,x,\mathbf{u},\mathbf{d}) \notin B) \geq P$, where $\mathbf{u}(t) = \pi(x(t))$.

Since $\mathbf{d}(t) = d \sim \mathscr{N}(\mu,\sigma^2)$, for compactness, throughout the rest of the paper whenever we refer to $\mathbf{d}$ we intend it in the form of Assumption 3, unless otherwise stated.

If we let $Z^1 = Z^2 = 1$, $H = -\delta$, $G^1(x^1) = x^1$, $g^1 = (St, v_T)^T$ and $N = 1$, the bad set of the application scenario, as depicted in Figure 1, can be written in the form of Assumption 4. We have proven in Proposition 1 that the flows of $x_f$ and $v_f$ are order preserving with respect to $\mathbf{u}$. Therefore, since $G^1(x^1) = x^1 = (x_f, v_f)^T$, Assumption 5 is also valid for $\Sigma_{app}$.

## III. SOLUTIONS

### A. Solution to Problem 1

Before proposing the solution, we define the $P$-safety capture set.

**Definition 6:** The $P$-safety capture set ($P \in (0,1)$) for a given control input signal $\mathbf{u} \in S(U)$ is defined as

$$C_{\mathbf{u}}(P) := \{x \in X \mid Pr(\forall t \in \mathbb{R}_+, \phi(t,x,\mathbf{u},\mathbf{d}) \notin B) < P\}.$$

The following Lemma shows that $C_{\mathbf{u}}(P)$ can be written as the union of two sets, which is convenient for computational purposes.

**Lemma 1:** The $P$-safety capture set for a given control input signal $\mathbf{u} \in S(U)$, for $B$ in the form of Assumption 4, can be written as $C_{\mathbf{u}} = \mathbf{S}_1^{\mathbf{u}} \cup \mathbf{S}_2^{\mathbf{u}}$, where

$$\mathbf{S}_1^{\mathbf{u}} := \left\{x \in X \;\middle|\; Pr\left(\forall t \in \mathbb{R}_+, Z^1 h^1(\phi^1(t,x^1,\mathbf{u})) - \\ Z^2 h^2(\phi^2(t,x^2,\mathbf{d})) \leq H\right) < P\right\},$$

$$\mathbf{S}_2^{\mathbf{u}} := \{x \in X \mid \exists\; t \in \mathbb{R}_+, \exists\; j \in \{1,...,N\} \; s.t. \\ G^j(\phi^1(t,x^1,\mathbf{u})) > g^j\}.$$

*Proof:* The bad set based on Assumption 4 is $B = B_1 \cup B_2$. According to Definition 6, the $P$-safety capture set for input signal $\mathbf{u}$ for this bad set is given by

$$C_{\mathbf{u}}(P) = \{x \in X \mid Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_1 \wedge \\ \phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2, \; \forall t \in \mathbb{R}_+) < P\}. \quad (3)$$

Let the set $S$ be defined as

$$S := \{x \in X \mid \exists\; t \in \mathbb{R}_+ \; s.t. \; \phi(t,x,\mathbf{u},\mathbf{d}) \in B_1\}, \quad (4)$$

which based on Assumption 4 is

$$\{x \in X \mid \exists\; t \in \mathbb{R}_+, \exists\; j \in \{1,...,N\} \; s.t. \; G^j(\phi^1(t,x^1,\mathbf{u})) > g^j\}. \quad (5)$$

We can write (3) in the form of $C_{\mathbf{u}}(P) = C_{\mathbf{u}}^S(P) \cup C_{\mathbf{u}}^{S^c}(P)$, where for $\mathbb{S} \in \{S, S^c\}$ we define

$$C_{\mathbf{u}}^{\mathbb{S}}(P) := \{x \in \mathbb{S} \mid Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_1 \wedge \\ \phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2, \forall t \in \mathbb{R}_+) < P\}. \quad (6)$$

If $x \in S$, since from Assumption 4 for all $j \in \{1,...,N\}$, $G^j$ is not function of the disturbance input $\mathbf{d}$, then from (4) and (5) we have $Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_1, \; \forall t \in \mathbb{R}_+) = Pr(\forall j \in \{1,...,N\}, \forall t \in \mathbb{R}_+, G^j(\phi^1(t,x^1,\mathbf{u})) \leq g^j) = 0$. Therefore, if $x \in S$ we have $Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_1 \wedge \phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2, \forall t \in \mathbb{R}_+) = 0 < P$, which is true for all $P \in (0,1)$. This implies that the capture set, from equations $C_{\mathbf{u}}(P) = C_{\mathbf{u}}^S(P) \cup C_{\mathbf{u}}^{S^c}(P)$ and (6), can be written as $C_{\mathbf{u}}(P) = S \cup C_{\mathbf{u}}^{S^c}(P)$.

If $x \in S^c$, similarly, from Assumption 4, equation (4) and equation (5), we obtain $Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_1,\ \forall t \in \mathbb{R}_+) = 1$, which is independent of the event $\phi(t,x,\mathbf{u},\mathbf{d}) \in B_2$. Therefore, if $x \in S^c$ we have

$$Pr\left(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_1 \wedge \phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2,\ \forall t \in \mathbb{R}_+\right) =$$

$$Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_1, \forall t \in \mathbb{R}_+) Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2, \forall t \in \mathbb{R}_+), \tag{7}$$

which since $Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_1, \forall t \in \mathbb{R}_+) = 1$, can be written in the form

$$Pr\left(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2,\ \forall t \in \mathbb{R}_+\right). \tag{8}$$

From equations (6), (7) and (8), the capture set can be written in the form $C_{\mathbf{u}}(P) = S \cup \{x \in S^c \mid Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2, \forall t \in \mathbb{R}_+) < P\}$. Since for any set $S$, $\{x \in S \ | Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2, \forall t \in \mathbb{R}_+) < P\} \subset S$, and $S \cup S^c = X$, then

$$C_{\mathbf{u}}(P) = S \cup \{x \in X \ | Pr\left(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B_2, \forall t \in \mathbb{R}_+\right) < P\}. \tag{9}$$

If we replace $S$ in (9) with its definition from (4) and (5), and use the definition of $B_2$ from Assumption 4, we can write (9) in the form of the statement of the Lemma. ■

Using the $Q$-notation [18], the following lemma provides a convenient way to compute set $\mathbf{S}_1^{\mathbf{u}}$ in Lemma 1.

**Lemma 2:** Let

$$F^{t,x,\mathbf{u}}(\mathbf{d}) := Z^1 h^1(\phi^1(t,x^1,\mathbf{u})) - Z^2 h^2(\phi^2(t,x^2,\mathbf{d})),$$

with $F_i^{t,x,\mathbf{u}}$ denoting the $i$th component of $F^{t,x,\mathbf{u}}$, and let $\bar{d} := \mu + \sigma Q^{-1}(P)$, and the pair $(t^*, i^*)$ (not necessarily unique) be such that

$$(t^*, i^*) = \arg \min_{\substack{t \in \mathbb{R}_+ \\ i \in \{1,...,r\}}} \left(H_i - F_i^{t,x,\mathbf{u}}(\bar{d})\right).$$

Then, we have

$$\mathbf{S}_1^{\mathbf{u}} = \left\{ x \in X \ \middle| \ H_{i^*} < F_{i^*}^{t^*,x,\mathbf{u}}(\bar{d}) \right\}.$$

*Proof:* (⇒) Since based on Assumption 2, the function $h^2(\phi^2(t,x^2,\mathbf{d}))$ is strictly order preserving with respect to $\mathbf{d}$, then based on Assumption 4, $Z^2 h^2(\phi^2(t,x^2,\mathbf{d}))$ is also strictly order preserving with respect to $\mathbf{d}$, and since $h^1(\phi^1(t,x^1,\mathbf{u}))$ is not a function of $\mathbf{d}$, then $F^{t,x,\mathbf{u}}(\mathbf{d})$ is a strictly decreasing function of $\mathbf{d}$, and therefore invertible. Hence, we can define

$$\left(F_i^{t,x,\mathbf{u}}\right)^{-1}(\alpha) := \left\{ \mathbf{d} \in \mathbb{R} \ \middle| \ F_i^{t,x,\mathbf{u}}(\mathbf{d}) = \alpha \right\}. \tag{10}$$

Using this property and the fact that in this part of the proof $x \in \mathbf{S}_1^{\mathbf{u}}$, based on the definition of $\mathbf{S}_1^{\mathbf{u}}$ from Lemma 1 we have

$$Pr\left(\forall t \in \mathbb{R}_+, Z^1 h^1(\phi^1(t,x^1,\mathbf{u})) - Z^2 h^2(\phi^2(t,x^2,\mathbf{d})) \leq H\right) =$$

$$Pr\left(\forall t \in \mathbb{R}_+, F^{t,x,\mathbf{u}}(\mathbf{d}) \leq H\right) =$$

$$Pr\left(\forall t \in \mathbb{R}_+, \forall i \in \{1,...,r\}, F_i^{t,x,\mathbf{u}}(\mathbf{d}) \leq H_i\right) =$$

$$Pr\left(\forall t \in \mathbb{R}_+, \forall i \in \{1,...,r\}, \mathbf{d} \geq (F_i^{t,x,\mathbf{u}})^{-1}(H_i)\right) =$$

$$Pr\left(\mathbf{d} \geq \max_{\substack{t \in \mathbb{R}_+ \\ i \in \{1,...,r\}}} (F_i^{t,x,\mathbf{u}})^{-1}(H_i)\right) =$$

$$\min_{\substack{t \in \mathbb{R}_+ \\ i \in \{1,...,r\}}} Pr\left(\mathbf{d} \geq \left(F_i^{t,x,\mathbf{u}}\right)^{-1}(H_i)\right) < P. \tag{11}$$

We prove that $H_{i^*} < F_{i^*}^{t^*,x,\mathbf{u}}(\bar{d})$. Assume that by contradiction $H_{i^*} \geq F_{i^*}^{t^*,x,\mathbf{u}}(\bar{d})$. Then based on the definition of the pair $(t^*, i^*)$, for all $t \in \mathbb{R}_+$ and $i \in \{1,...,r\}$, we must have $H_i \geq F_i^{t,x,\mathbf{u}}(\bar{d})$, which because of the strictly decreasing property of $F_i^{t,x,\mathbf{u}}(\mathbf{d})$ with respect to $\mathbf{d}$, can be written as

$$\forall t \in \mathbb{R}_+,\ \forall i \in \{1,...,r\},\ (F_i^{t,x,\mathbf{u}})^{-1}(H_i) \leq \bar{d}. \tag{12}$$

Since based on Assumption 3, for all $t \in \mathbb{R}_+$ we have $\mathbf{d}(t) = d \sim \mathcal{N}(\mu, \sigma^2)$, then using the $Q$-notation and the definition of $\bar{d}$ we have $Pr(\mathbf{d} \geq \bar{d}) = P$. This relationship along with equation (12) implies that $Pr(\mathbf{d} \geq (F_i^{t,x,\mathbf{u}})^{-1}(H_i)) \geq P$, for all $t \in \mathbb{R}_+$ and $i \in \{1,...,r\}$. This contradicts (11). Therefore, $H_{i^*} < F_{i^*}^{t^*,x,\mathbf{u}}(\bar{d})$.

(⇐) If for $x \in X$ we have $H_{i^*} < F_{i^*}^{t^*,x,\mathbf{u}}(\bar{d})$, then because of the strictly decreasing property of $F_{i^*}^{t^*,x,\mathbf{u}}$ with respect to $\mathbf{d}$, we have $(F_{i^*}^{t^*,x,\mathbf{u}})^{-1}(H_{i^*}) > \bar{d}$. This relationship along with $Pr(\mathbf{d} \geq \bar{d}) = P$ implies that $Pr(\mathbf{d} \geq (F_{i^*}^{t^*,x,\mathbf{u}})^{-1}(H_{i^*})) < P$, which because of the strictly decreasing property of $F_{i^*}^{t^*,x,\mathbf{u}}$ with respect to $\mathbf{d}$, can be written as $Pr(H_{i^*} \geq F_{i^*}^{t^*,x,\mathbf{u}}(\mathbf{d})) < P$. Since $Pr(\forall t \in \mathbb{R}_+,\ \forall i \in \{1,...,r\},\ H_i \geq F_i^{t,x,\mathbf{u}}(\mathbf{d})) \leq Pr(H_{i^*} \geq F_{i^*}^{t^*,x,\mathbf{u}}(\mathbf{d}))$, then we have

$$Pr(\forall t \in \mathbb{R}_+,\ \forall i \in \{1,...,r\},\ F_i^{t,x,\mathbf{u}}(\mathbf{d}) \leq H_i) =$$

$$Pr(\forall t \in \mathbb{R}_+,\ F^{t,x,\mathbf{u}}(\mathbf{d}) \leq H) < P. \tag{13}$$

If we replace $F^{t,x,\mathbf{u}}(\mathbf{d})$ in (13) with its definition from the statement of the lemma, we conclude $x \in \mathbf{S}_1^{\mathbf{u}}$. ■

Lemmas 1 and 2 together provide a means to compute the capture set. These lemmas are used to prove the following theorem and in Section IV to provide algorithmic procedures for the computation of the capture set.

**Theorem 1:** For system $\Sigma^*$, with Assumptions 1-5, $x \in \mathscr{W}$ if and only if $x \notin C_{u_m}(P)$.

*Proof:* (⇐) If $x \notin C_{u_m}(P)$ then $x \in C_{u_m}^c(P)$. Therefore, $Pr(\phi(t,x,u_m,\mathbf{d}) \notin B,\ \forall t \in \mathbb{R}_+) \geq P$, which implies $x \in \mathscr{W}$.

(⇒) If $x \in \mathscr{W}$, then there is control input signal $\mathbf{u}' \in S(U)$ such that $Pr(\phi(t,x,\mathbf{u}',\mathbf{d}) \notin B, \forall t \in \mathbb{R}_+) \geq P$. If we replace the relation "<" in Definition 6 with "≥", and use the results of Lemmas 1 and 2, we conclude that for $x \in \mathscr{W}$ we must have that there is control input signal $\mathbf{u}' \in S(U)$ such that

$$Pr(\forall t \in \mathbb{R}_+, Z^1 h^1(\phi^1(t,x^1,\mathbf{u}')) - Z^2 h^2(\phi^2(t,x^2,\mathbf{d})) \leq H) \geq P$$

$$\text{and } \forall t \in \mathbb{R}_+,\ \forall j \in \{1,...,N\},\ G^j(\phi^1(t,x^1,\mathbf{u}')) \leq g^j. \tag{14}$$

We prove that $x \notin C_{u_m}(P)$. Assume that by contradiction $x \in C_{u_m}(P)$, then based on Lemmas 1 and 2, at least one of the following cases must hold. Case (1): $H_{i^*} < F_{i^*}^{t^*,x,u_m}(\bar{d})$, where

$$(t^*, i^*) = \arg \min_{\substack{t \in \mathbb{R}_+ \\ i \in \{1,...,r\}}} \left(H_i - F_i^{t,x,u_m}(\bar{d})\right); \tag{15}$$

or Case (2): there is a time $t \in \mathbb{R}_+$ and a $j \in \{1,...,N\}$ such that $G^j(\phi^1(t,x^1,u_m)) > g^j$.

Case (1): If $x \in \mathscr{W}$, then according to (14) and Lemma

2 (with relation "$<$" replaced with "$\geq$"), there is a control input signal $\mathbf{u}' \in S(U)$ such that $H_{i'} \geq F_{i'}^{t',x,\mathbf{u}'}(\bar{d})$, where

$$(t',i') = \arg \min_{\substack{t\in\mathbb{R}_+\\ i\in\{1,\ldots,r\}}} \left(H_i - F_i^{t,x,\mathbf{u}'}(\bar{d})\right). \tag{16}$$

If $H_{i'} \geq F_{i'}^{t',x,\mathbf{u}'}(\bar{d})$, then according to (16) we also have $H_{i^*} \geq F_{i^*}^{t^*,x,\mathbf{u}'}(\bar{d})$. This result along with equation $H_{i^*} < F_{i^*}^{t^*,x,u_m}(\bar{d})$, which is the main assumption in Case (1), implies that $F_{i^*}^{t^*,x,\mathbf{u}'}(\bar{d}) \leq H_{i^*} < F_{i^*}^{t^*,x,u_m}(\bar{d})$, which based on the definition of $F_{i^*}^{t^*,x,u_m}$ and $F_{i^*}^{t^*,x,\mathbf{u}'}$ from Lemma 2 can be expanded to

$$\sum_{l_1=1}^{\dim(O^1)} Z^1_{i^*l_1} h^1_{l_1}(\phi^1(t^*,x^1,\mathbf{u}')) - \sum_{l_2=1}^{\dim(O^2)} Z^2_{i^*l_2} h^2_{l_2}(\phi^2(t^*,x^2,\bar{d})) <$$

$$\sum_{l_3=1}^{\dim(O^1)} Z^1_{i^*l_3} h^1_{l_3}(\phi^1(t^*,x^1,u_m)) - \sum_{l_4=1}^{\dim(O^2)} Z^2_{i^*l_4} h^2_{l_4}(\phi^2(t^*,x^2,\bar{d}))$$

$$\Rightarrow \sum_{l=1}^{\dim(O^1)} Z^1_{i^*l}\left[h^1_l(\phi^1(t^*,x^1,\mathbf{u}')) - h^1_l(\phi^1(t^*,x^1,u_m))\right] < 0. \tag{17}$$

Since $u_m$ is the minimal control input and based on Assumption 1, $h^1$ is an order preserving function of $\mathbf{u}$, then for all $l \in \{1,...,\dim(O^1)\}$ we have $h^1_l(\phi^1(t^*,x^1,\mathbf{u}')) - h^1_l(\phi^1(t^*,x^1,u_m)) \geq 0$. In turn, from Assumption 4 we have $Z^1_{i^*l} \geq 0$. These two statements together contradict (17). Therefore, $H_{i^*} \geq F_{i^*}^{t^*,x,u_m}(\bar{d})$ and Case (1) cannot hold.

Case (2): If $x \in C_{u_m}(P)$, then based on Lemma 1 we must have a time $\tau \in \mathbb{R}_+$ and a $j \in \{1,...,N\}$ such that $G^j(\phi^1(\tau,x^1,u_m)) > g^j$. Because of the order preserving property of function $G^j(\phi^1(\tau,x^1,\mathbf{u}))$ with respect to $\mathbf{u}$ based on Assumption 5, for all $\mathbf{u} \in S(U)$ we have $G^j(\phi^1(\tau,x^1,u_m)) \leq G^j(\phi^1(\tau,x^1,\mathbf{u}))$. Therefore, if $G^j(\phi^1(\tau,x^1,u_m)) > g^j$, then we also have $G^j(\phi^1(\tau,x^1,\mathbf{u})) > g^j$ for all $\mathbf{u} \in S(U)$. Since $x \in \mathscr{W}$, then based on (14) there is also a control input signal $\mathbf{u}' \in S(U)$ such that for all $t \in \mathbb{R}_+$ and $j \in \{1,...,N\}$, $G^j(\phi^1(t,x^1,\mathbf{u}')) \leq g^j$. Since this statement is valid for all $t \in \mathbb{R}_+$, then $G^j(\phi^1(\tau,x^1,\mathbf{u}') \leq g^j$, which contradicts our previous statement that for all $\mathbf{u} \in S(U)$ we have $G^j(\phi^1(\tau,x^1,\mathbf{u})) > g^j$. Therefore, there is no $j \in \{1,...,N\}$ and $\tau \in \mathbb{R}_+$ such that $G^j(\phi^1(\tau,x^1,u_m)) > g^j$. This implies that Case (2) cannot hold. Since neither of Case 1 or Case 2 holds, then $x \notin C_{u_m}(P)$. ∎

### B. Solution to Problem 2

We define the boundary of a set $C$ as $\partial C = Cl(C) \cap Cl(C^c)$, where $Cl(C)$ represents the closure of the set $C$. We consider the feedback control map

$$\pi(x) = \begin{cases} U & \text{if } x \notin \{C_{u_m}(P) \cup \partial C_{u_m}(P)\} \\ u_m & \text{if } x \in \{C_{u_m}(P) \cup \partial C_{u_m}(P)\} \end{cases}, \tag{18}$$

and state the following theorem.

**Theorem 2:** For system $\Sigma^*$, with Assumptions 1-5, for all $x \in \mathscr{W}$ the feedback map $\pi : X \to U$, as defined in (18), guarantees that $Pr(\phi(t,x,\mathbf{u},\mathbf{d}) \notin B, \forall t \in \mathbb{R}_+) \geq P$.

*Proof:* Recall that $\mathbf{d}(t) \equiv d$ for some $d \sim \mathscr{N}(\mu,\sigma^2)$. Setting $\bar{d} = \mu + \sigma Q^{-1}(P)$, this implies that $\Pr(\mathbf{d}(t) \geq \bar{d}, \forall t \in \mathbb{R}_+) = P$. Consequently, it suffices to show that

$$\phi(t,x,\mathbf{u},d) \notin B, \quad \forall t \in \mathbb{R}_+, \quad \forall d \geq \bar{d}.$$

We do this with a contradiction argument. Assume therefore that there exist $d \geq \bar{d}$ and $t^* \in \mathbb{R}_+$ such that

$$\phi(t^*,x,\mathbf{u},d) \in B. \tag{19}$$

As $B \subset C_{u_m}(P)$, it is clear that $\phi(t^*,x,\mathbf{u},d) \in C_{u_m}(P)$. Defining $\bar{t} = \sup\{t \in [0,t^*] \mid \phi(t,x,\mathbf{u},d) \in C_{u_m}(P)^c\}$, it follows from the continuity of the flow with respect to time that $\bar{x} = \phi(\bar{t},x,\mathbf{u},d) \in \partial C_{u_m}(P)$. Moreover, by the very definition of $\bar{t}$ and the static feedback controller $\pi$, $\mathbf{u}(t) = u_m$ for all $t \geq \bar{t}$. Next notice that $C_{u_m}(P)$ is open. Indeed this is a consequence of Lemma 1 and Lemma 2. Openness of $C_{u_m}(P)$ together with these Lemmas implies in turn that for all $j \in \{1,\ldots,N\}$ and $t \in \mathbb{R}_+$, there exist $k \in \{1,\ldots,r\}$ and $i \in \{1,\ldots,q\}$ such that

$$F_k^{t,\bar{x},u_m}(\bar{d}) \leq H_k \quad \text{and} \quad G_i^j(\phi^1(t,\bar{x}^1,u_m)) \leq g_i^j. \tag{20}$$

Finally, by the order preserving property of $d \to \phi^2(t,x,d)$, this implies also that

$$F_k^{t,\bar{x},u_m}(d) \leq H_k, \quad \forall t \in \mathbb{R}_+. \tag{21}$$

However, (20)-(21) assure that $\phi(t,x,\mathbf{u},d) \notin B$ for all $t \geq \bar{t}$ contradicting (19). ∎

## IV. IMPLEMENTATION AND SIMULATION

### A. Algorithm

Based on Theorem 2, if we can calculate $C_{u_m}(P)$, then equation (18) provides a control map that guarantees the minimum safety $P$. Also, Lemma 1 and Lemma 2 together provide us with a relationship that can be used to compute the capture set. For $\mathbf{u} = u_m$ we have

$$C_{u_m}(P) = \mathbf{S}_1^{u_m} \cup \mathbf{S}_2^{u_m}, \tag{22}$$

where $\mathbf{S}_1^{u_m}$ and $\mathbf{S}_2^{u_m}$ can be computed from Lemma 2 and Lemma 1, respectively.

Equation (22) requires to determine $\min_{t\in\mathbb{R}_+}(H_i - F_i^{t,x,u_m}(\bar{d}))$ and to check whether there is a $t \in \mathbb{R}_+$ such that $G^j(\phi^1(t,x^1,u_m)) > g^j$. The following assumption allows us to determine such a minimum and to check the existence of such $t$ on a bounded time interval $(0,\tau]$ with known $\tau$.

**Assumption 6:** Let $\mathrm{u}^1 = u$ and $\mathrm{u}^2 = d$, then $f^1(x^1,u)$ and $f^2(x^2,d)$ in system $\Sigma^*$ are in the following form:

$$\text{for } i \in \{1,2\},\ f^i(x^i,\mathrm{u}^i) = \begin{cases} \bar{f}^i(x^i,\mathrm{u}^i) & \text{if } \frac{d}{dt}h^i(x^i) > 0 \\ 0 & \text{if } \frac{d}{dt}h^i(x^i) \leq 0 \end{cases},$$

and there is a finite time $\tau \in (0,\infty)$ such that $\tau = \min\{t \in \mathbb{R}_+ \mid \frac{d}{dt}h^1(\phi^1(t,x^1,u_m)) = 0\}$.

Note that functions $f^1(x^1,u)$ and $f^2(x^2,d)$ of $\Sigma_{app}$ are in the above form. Moreover, we can set a $u_m$ such that for all $t \in \mathbb{R}_+$, $u_m - Dv_f^2(t) - a_r - a_s < 0$. Since $\frac{d}{dt}h^1(\phi^1(t,x^1,u_m)) = v_f(t)$ and $\dot{v}_f(t) = u_m - Dv_f^2(t) - a_r - a_s$, we can guarantee that $\tau$ is a finite time, which physically relates to the fact that a continuous braking effort will make the velocity of FV reach zero in a finite time.

**Proposition 3:** Assumption 6 leads to the following relationship

$$\min_{\substack{t\in\mathbb{R}_+\\ i\in\{1,\ldots,r\}}} \left(H_i - F_i^{t,x,u_m}(\bar{d})\right) = \min_{\substack{t\in(0,\tau]\\ i\in\{1,\ldots,r\}}} \left(H_i - F_i^{t,x,u_m}(\bar{d})\right). \quad (23)$$

Also, for any $j \in \{1,\ldots,N\}$, there is a $t \in \mathbb{R}_+$ such that $G^j(\phi^1(t,x^1,u_m)) > g^j$, if and only if there is a $t \in (0,\tau]$ such that $G^j(\phi^1(t,x^1,u_m)) > g^j$.

*Proof:* In the first part, we prove that equation (23) holds. Based on the model of $\Sigma^*$ and Assumption 6, for all $t > \tau$ we have $\dot{x}^1(t) = \dot{\phi}^1(t,x^1,u_m) = 0$. Therefore,

$$Z^1\frac{d}{dt}h^1(x^1(t)) - Z^2\frac{d}{dt}h^2(x^2(t)) = Z^1 J_{h^1}(x^1(t))\dot{x}^1(t) - \\ Z^2\frac{d}{dt}h^2(x^2(t)) = -Z^2\frac{d}{dt}h^2(x^2(t)), \quad (24)$$

where $J_{h^1}(x^1(t))$ is the Jacobian of function $h^1$ defined as $J_{h^1}(x^1(t))_{ij} = \frac{\partial h^1_i}{\partial x^1_j}$. Since $\frac{d}{dt}h^2(x^2(t)) = J_{h^2}(x^2(t))\dot{x}^2(t)$, then based on Assumption 6 we have that $\frac{d}{dt}h^2(x^2(t)) \geq 0$. Based on (24), the definition of $F^{t,x,u_m}(\bar{d})$, and the fact that $Z^2$ is non-negative, this implies that for all $t \geq \tau$ we have $\frac{d}{dt}(H - F^{t,x,u_m}(\bar{d})) = -\frac{d}{dt}(F^{t,x,u_m}(\bar{d})) = Z^2\frac{d}{dt}h^2(x^2(t)) \geq 0$. Therefore, $\arg\min_{t\in\mathbb{R}_+}(H - F^{t,x,u_m}(\bar{d})) \in (0,\tau]$, which implies that (23) holds. In the next part, we prove that for a $j \in \{1,\ldots,N\}$ there is a $t \in \mathbb{R}_+$ such that $G^j(\phi^1(t,x^1,u_m)) > g^j$, if and only if there is a $t \in (0,\tau]$ such that $G^j(\phi^1(t,x^1,u_m)) > g^j$.

($\Rightarrow$) We assume that there is a $t \in \mathbb{R}_+$ and $j \in \{1,\ldots,N\}$ such that $G^j(\phi^1(t,x^1,u_m)) > g^j$, and by contradiction $t \notin (0,\tau]$. This implies that we must have for all $t \in (0,\tau]$ and $j \in \{1,\ldots,N\}$, $G^j(x^1(t)) \leq g^j$, while there is a $t^* \in (\tau,\infty)$ such that $G^j(x^1(t^*)) > g^j$. Therefore, $G^j(x^1(t^*)) > \max_{t\in(0,\tau]} G^j(x^1(t))$. For all $t \geq \tau$ we have $\frac{d}{dt}G^j(x^1(t)) = J_{G^j}(x^1(t))\dot{x}^1 = 0$, where $J_{G^j}(x^1(t))$ is the Jacobian of function $G^j$. This implies that $G^j(x^1(t^*)) = G^j(x^1(\tau))$, which contradicts our previous statement that $G^j(x^1(t^*)) > \max_{t\in(0,\tau]} G^j(x^1(t))$. Therefore, $t \in (0,\tau]$.

($\Leftarrow$) If there is a time $t \in (0,\tau]$ and a $j \in \{1,\ldots,N\}$ such that $G^j(x^1(t)) > g^j$, since $(0,\tau] \subset (0,\infty)$, the relationship $t \in (0,\infty)$ is trivially satisfied. ■

In order to implement the feedback map (18), we use a discrete-time algorithm. We use the forward Euler approximation for discretization. The time step size is denoted by $\Delta t$. We also denote the state of the system at step $k$ by $x[k] = (x^1[k], x^2[k])^T$. Therefore, $x(k\Delta t) = x[k]$, where $x(k\Delta t)$ is the state of the system at time $k\Delta t$ in the continuous-time model. We also define the function $\mathbf{F}_i[k]$ as follows.

$$\mathbf{F}_i[k] := \sum_{l_1=1}^{\dim(O^1)} Z^1_{il_1} h^1_{l_1}(x^1[k]) - \sum_{l_2=1}^{\dim(O^2)} Z^2_{il_2} h^2_{l_2}(x^2[k]). \quad (25)$$

All other notations are similar to the continuous-time model. Note that $x^1[k]$ and $x^2[k]$ in (25) depend on the initial condition and the inputs. Algorithm 1 shows the discrete-time implementation of (18).

**Algorithm 1** Control Feedback Computation

```
Require: $x[0] = (x^1[0], x^2[0])^T$: the current state.
  $u[0]$: the current input to the system.
  $\bar{d} \leftarrow \mu + \sigma Q^{-1}(P)$ (use z-table for $Q^{-1}(P)$ [16])
  $k \leftarrow 0$
  $x[k+1] \leftarrow x[k] + \Delta t(f(x[k], u[0], \bar{d}))$
  while $\frac{d}{dt}h^1(x^1[k]) = \frac{h^1(x^1[k+1]) - h^1(x^1[k])}{\Delta t} > 0$ do
    for $i = 1$ to $i = r$ do
      if $H_i - \mathbf{F}_i[k] \leq 0$ then
        $u \leftarrow u_m$, STOP.
      end if
    end for
    for $j = 1$ to $j = N$ do
      if $G^j(x^1[k]) - g^j \geq 0$ then
        $u \leftarrow u_m$, STOP.
      end if
    end for
    $k \leftarrow k + 1$
    $x[k+1] = x[k] + \Delta t(f(x[k], u_m, \bar{d}))$
  end while
  $u \in U$
```

### *B. Simulations and Data Analysis*

Figure 2 shows the path that was used to generate the data for both identifying the parameters of PV model, that is, $a$, $b$, $\mu$ and $\sigma$ in equation (2b) (note that in equation (2b), $d \sim \mathcal{N}(\mu, \sigma^2)$), and validating our algorithm. This path is located in Ann Arbor, Michigan, and it is 11 *km* long and consists of 30 study areas. The study area is a part of the road at which the driver frequently reduces his/her speed such as intersections, roundabouts or stop signs. The data of any of these regions is used to identify the model parameters. At each study area, the vehicle software provides the relative distance to the study area's lowest velocity point $St - x_f$ (e.g., the stop sign at an intersection), and the vehicle's lowest velocity $v_T$. Position and speed of our test vehicle at study areas are depicted in Figure 3.

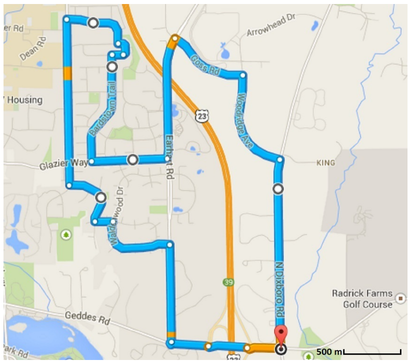


Fig. 2: The path that is used for data acquisition.

We have used the least squares method to calculate the parameters $a$, $b$, $\mu$ and $\sigma$ from the data of 420 trajectories, as depicted in Figure 3. In particular, using equations (1) and

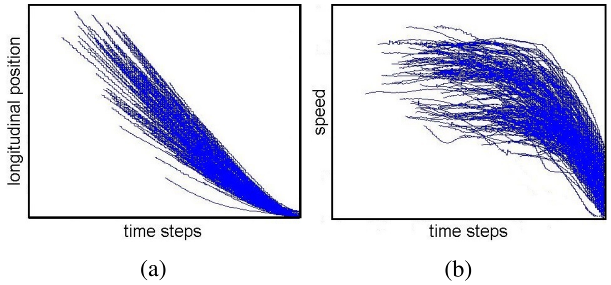


Fig. 3: Trajectories of position and speed of 420 profiles of the test vehicle at study areas [units and numbers are removed as they are proprietary information].

(2b) for $v_p[k] \geq 0$ we have

$$\begin{bmatrix} x_p[k+1] \\ v_p[k+1] \end{bmatrix} = \begin{bmatrix} x_p[k] + \Delta t v_p[k] \\ v_p[k] + \Delta t(a x_p[k] + b v_p[k] + \mu) \end{bmatrix}. \quad (26)$$

By replacing $x_p[k]$ in the second equation of (26) with $x_p[k-1] + \Delta t v_p[k-1]$, we obtain $v_p[k+1] = a(\Delta t x_p[k-1] + \Delta t^2 v_p[k-1]) + (1+b\Delta t)v_p[k] + \mu\Delta t$. We define the new parameters $a' = a$, $b' = 1 + b\Delta t$ and $\mu' = \mu$. Minimizing the mean square error for speed leads to the following optimization problem:

$$\min_X ||CX - D||^2, \text{ with } X = (a', b', \mu')^T, \quad (27)$$

where $C(1,1) = \Delta t x_p[1]$, $C(1,2) = v_p[1]$, $C(1,3) = \Delta t$, and for $k \geq 2$ we have $C(k,1) = \Delta t x_p[k-1] + \Delta t^2 v_p[k-1]$, $C(k,2) = v_p[k]$, $C(k,3) = \Delta t$, $D(k-1) = v_p[k]$, and the variance can be computed using

$$\sigma^2 = \frac{\sum_{i=1}^{N_d} \left| a x_p[i] + b v_p[i] + \mu - a_p[i] \right|^2}{N_d},$$

where $N_d$ is the number of data points.

In order to test that Algorithm 1 can save the vehicles from collisions $100P\%$ of the total collision instances, we determined the empirical safety level as follows. By the law of large numbers, for an event $x$ with mean $\mu$ we have $Pr(\lim_{T\to\infty} \frac{N}{T} = \mu) = 1$, where $\frac{N}{T}$ is the fraction of times in $T$ trials that the event $x$ has been observed. We then generated an initial condition for FV and chose a trajectory for PV among our available data, which consists of 420 trajectories, randomly, for $T$ times. Based on the law of large numbers, for a large $T$ we expect to observe approximately $(1-P)T$ number of collisions. The logic diagram of our tests is shown in Figure 4. The result of running the algorithm for $T = $10000 times is shown in Table I, where we can see the empirical safety level is very close to the safety level that the feedback map (18) provides, validating our supervisor design. In order to verify that the parameters are not overfitted, we must check how a model that is constructed based on a limited set of data will respond to a new dataset. We use the k-fold cross validation method, as introduced in [17], with $k = 10$. Thus, we partition our available data of 420 trajectories, depicted in Figure 3, into 10 groups (each group with 42 trajectories) and solve the minimization problem (27) using the data of 9 groups and run the tests of Figure 4 on the 10th group. We repeat this for all 10 groups and compare their average empirical safety level with the expected safety level as we have done in Table I. The result is shown in Table II. Table II shows that the model parameters $a$, $b$, $\mu$ and $\sigma$ are not overfitted, since the algorithm leads to an empirical safety level $1 - \frac{N}{T}$ close to $P$ independent of the training data.

| Safety ($100P\%$) | Tests ($T$) | Empirical safety $(100(1-\frac{N}{T}))$ |
|---|---|---|
| 70% | 10000 | 72.2% |
| 80% | 10000 | 82.3% |
| 90% | 10000 | 91.7% |

TABLE I: Result of running the test of Figure 4 for $P = 0.7$, $P = 0.8$ and $P = 0.9$, each for 10000 times.

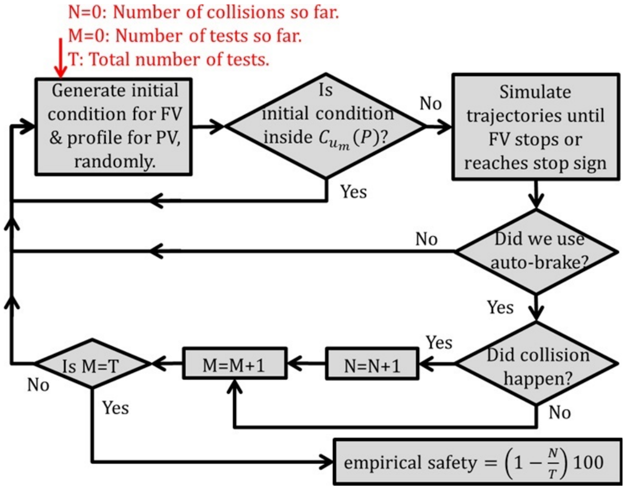

N=0: Number of collisions so far.
M=0: Number of tests so far.
T: Total number of tests.
Generate initial condition for FV & profile for PV, randomly.
Is initial condition inside $C_{u_m}(P)$?
No
Yes
Simulate trajectories until FV stops or reaches stop sign
Did we use auto-brake?
No
Yes
Did collision happen?
Yes
No
N=N+1
M=M+1
Is M=T
No
Yes
empirical safety = $\left(1-\frac{N}{T}\right)100$

Fig. 4: Test to evaluate the safety of the system.

| Safety ($100P\%$) | 70% | 80% | 90% |
|---|---|---|---|
| Group 1 | 67.8% | 80.3% | 88.9% |
| Group 2 | 69% | 79.8% | 90.3% |
| Group 3 | 68.8% | 80.1% | 89.2% |
| Group 4 | 72.1% | 81.3% | 92.1% |
| Group 5 | 71.3% | 80.2% | 91.1% |
| Group 6 | 69.8% | 78.3% | 90.5% |
| Group 7 | 71.1% | 80.3% | 93% |
| Group 8 | 68.3% | 79.5% | 92.1% |
| Group 9 | 70.4% | 80% | 91.2% |
| Group 10 | 70.3% | 80.3% | 88.9% |
| Average | 69.89% | 80.1% | 90.73% |

TABLE II: Result of 10-fold cross validation.

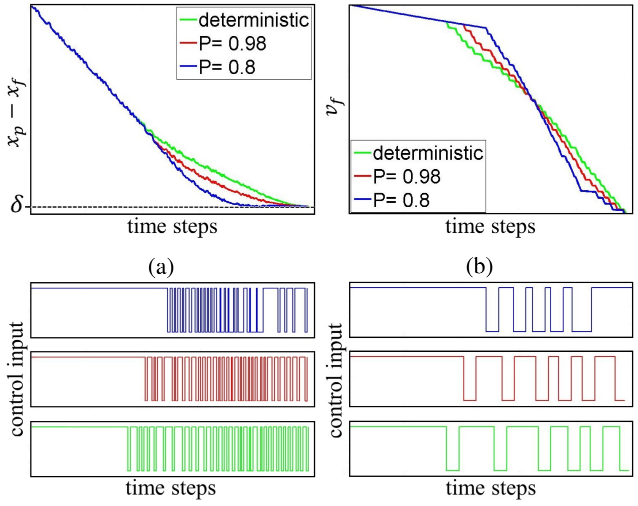
Since we have identified the parameters based on the data of the test vehicle, we have to verify that this model guarantees the minimum $P$-safety when the measurements of position and speed of PV are obtained through radar, as vehicles appear randomly in front of FV (the test vehicle). In Table III, we show the results. In this case, 68 trajectories of PV have been used. Table III shows that although we have identified the parameters based on the data of the test vehicle, they still can be used to ensure the desired safety level in the presence of the randomly appearing PVs detected from radar.

In Figure 5, we show the results of simulations for an

| Safety ($100P\%$) | Tests ($T$) | Empirical safety ($100(1-\frac{N}{T})$) |
|---|---|---|
| 70% | 5000 | 70.2% |
| 80% | 5000 | 79.8% |
| 90% | 5000 | 91.1% |

TABLE III: Results of running the test of Figure 4 for data of PV with the model built based on the data of the test vehicle.

arbitrary trajectory of PV chosen from our data for $P = 0.8$, $P = 0.98$ and the deterministic model as was suggested in [4] and [5]. Figure (5-a) shows the relative distance between vehicles, and Figure (5-b) shows $v_f$. In Figures (5-c) and (5-d), we have plotted the control input. In Figure (5-d), we have used a counter that keeps the control input $u_m$ on for at least $1s$ whenever the system exits $C_{u_m}$. This way, as we can see from the plots, the number of switches between $u_m$ and $u \in U$ has reduced significantly in Figure (5-d) compared to Figure (5-c). From all plots, it is clear that the deterministic model is more conservative than the stochastic one, and $P = 0.98$ is more conservative than $P = 0.8$.



(a)

(b)

(c)

(d)

Fig. 5: Simulation results [units and numbers are removed as they are proprietary information].

## V. CONCLUSIONS

In this work, we have provided a control feedback map that guarantees the safety of a class of stochastic order preserving systems with the minimum probability $P$. The results of Table I show that the safety probability is actually very close to $P$, indicating that the feedback map is not conservative. Table II shows that the parameters are not overfitted, and Table III further indicates that the parameters can be safely identified through data of a test vehicle and still capture behaviors of PV as obtained through radar.

Algorithm 1 along with an algorithm that provides both warnings to drivers and automatic brake, is eventually being implemented on a Prius vehicle. As part of the future works, we will investigate different models for the acceleration of PV instead of the linear function, in order to provide even more accurate models.

## REFERENCES

[1] C. Tomlin, J. Lygeros, S. S. Sastry, ”Controller synthesis for hybrid systems: The Hamilton-Jacobi approach,”, in *AAAI Spring Symposia*, 1999, pp. 192-197.

[2] C. Tomlin, J. Lygeros, S. S. Sastry, “A game theoretic approach to controller design for hybrid systems”, *Proceedings of IEEE*, vol. 88, pp. 949-970, 2000.

[3] C. Tomlin, I. Mitchell, A. M. Bayen, M. Oishi, “Computational techniques for the verification of hybrid systems”, *Proceeding of IEEE*, vol. 91, pp. 986-1001, 2003.

[4] R. Verma, D. Del Vecchio, “Semiautonomous multivehicle safety: A hybrid control approach”, *IEEE Robotics and Automation Magazine*, vol. 18, pp. 44-54, 2011.

[5] M. Hafner, D. Cunningham, L. Caminiti, and D. Del Vecchio, ”Cooperative collision avoidance at intersections: Algorithms and experiments”, *IEEE Trans. Intelligent Transportation Systems*, vol. 14, pp. 1162-1175, 2013.

[6] J. Duperret, M. Hafner, D. Del Vecchio, “Formal design of a provably safe robotic roundabout system”, in *Proc. IEEE/RSJ Conference on Intelligent Robots and Systems*, 2010, pp. 2006-2011.

[7] V. Desaraju, H. C. Ro, M. Yang, E. Tay, S. Roth, D. Del Vecchio, “Partial order techniques for vehicle collision avoidance: Application to an autonomous roundabout test-bed”, in *Proc. IEEE Conference on Robotics and Automation*, 2009, pp. 82-87.

[8] A. Colombo, D. Del Vecchio, “Efficient algorithms for collision avoidance at intersections”, in *Proc. of the ACM International Conference on Hybrid Systems: Computation and Control*, 2012, pp. 145-154.

[9] S. Prajna, A. Jadbabaie, G. J. Pappas, “A framework for worst-case and stochastic safety verification using barrier certificates”, *IEEE Trans. on Automatic Control*, vol. 52, pp. 1415-1428, 2007.

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

[11] J. Ding, A. Abate, C. Tomlin, “Optimal control of partially observable discrete time stochastic hybrid systems for safety specifications”, in *Proc. of American Control Conference*, 2013, pp. 6231-6236.

[12] “2012 motor vehicle crashes: Overview. U.S. Department of Transportation National Highway Safety Administration”, http://www-nrd.nhtsa.dot.gov/Pubs/811856.pdf

[13] “Intersection Safety”, http://safety.fhwa.dot.gov/intersection/

[14] G.S. Aoude, V.R. Desaraju, L.H. Stephens, 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.

[15] K. Hayashi, Y. Kojima, K. Abe, K. Oguri, “Prediction of stopping maneuver considering driver’s state”, in *Proc. of IEEE Conference on Intelligent Transportation Systems*, 2006, pp. 1191-1196.

[16] D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to Probability*, Second Edition, Athena Scientific, 2008.

[17] P. A. Devijver, J. Kittler, *Pattern Recognition: A Statistical Approach*, Prentice-Hall, 1982.

[18] G. K. Karagiannidis, A. S. Lioumpas, “An improved approximation for the Gaussian Q-Function”, *IEEE Communication Letters*, vol. 11, pp. 644-646, 2007.

[19] M. Forghani, D. Del Vecchio, “Order preserving properties of vehicle dynamics with respect to the driver’s input”, Technical Report, MIT, September 2014, http://hdl.handle.net/1721.1/90250.

[20] P. M. Esfahani, D. Chatterjee, J. Lygeros, “On a problem of stochastic reach-avoid set characterization”, in *Proc. IEEE Conference on Decision and Control*, 2011, pp. 7069-7074.