Received 29 March 2022; revised 5 July 2022; accepted 19 August 2022; Date of publication TBD; date of current version 2 September 2022.

*Digital Object Identifier 10.1109/OJCSYS.2022.TBD*

# Efficient Learning of Hyperrectangular Invariant Sets using Gaussian Processes

**Michael Enqi Cao[1] (Student Member, IEEE), Matthieu Bloch[1] (Senior Member, IEEE), Samuel Coogan[1,2] (Senior Member, IEEE)**

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, 30332, USA
[2]School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, 30332, USA

CORRESPONDING AUTHOR: Michael Enqi Cao (e-mail: mcao34@gatech.edu)

This project was partially supported by the National Science Foundation under awards #1749357 and #1836932, the Ford Motor Company, and Cisco Systems.

**ABSTRACT** We present a method for efficiently computing reachable sets and forward invariant sets for continuous-time systems with dynamics that include unknown components. Our main assumption is that, given any hyperrectangle of states, lower and upper bounds for the unknown components are available. With this assumption, the theory of mixed monotone systems allows us to formulate an efficient method for computing a hyperrectangular set that over-approximates the reachable set of the system. We then show a related approach that leads to sufficient conditions for identifying hyperrectangular sets that are forward invariant for the dynamics. We additionally show that set estimates tighten as the bounds on the unknown behavior tighten. Finally, we derive a method for satisfying our main assumption by modeling the unknown components as state-dependent Gaussian processes, providing bounds that are correct with high probability. A key benefit of our approach is to enable tractable computations for systems up to moderately high dimension that are subject to low dimensional uncertainty modeled as Gaussian processes, a class of systems that often appears in practice. We demonstrate our results on several examples, including a case study of a planar multirotor aerial vehicle.

**INDEX TERMS** Autonomous Systems, Safe Learning for Control, Stability of Nonlinear Systems

## I. Introduction

The calculation of reachable or forward invariant sets for a dynamical system is often a key component of safety verification. However, these calculations tend to suffer from the curse of dimensionality, and the complexity can be compounded if the true dynamics of the system are not fully known due to inaccuracies in the model or the presence of external disturbances. *Mixed monotone systems* theory has recently proved effective for efficiently estimating rectangular forward invariant sets and overapproximations of reachable sets [1], with applications to control of practical systems of around ten state dimensions [2], [3]. Further, this theory is able to accommodate unknown but bounded disturbances into this calculation [4]. We extend these ideas by leveraging Gaussian Process (GP) theory to efficiently calculate high-confidence bounds on unknown components of the dynamics in order to compute, with high probability, reachable sets and invariant sets. We then show that set estimates are improved as the bounds on the unknown components tighten due to, *e.g.*, measurements of the GP, which leads to several applications of this formulation in robust control and safety verification.

An $n$–dimensional system is mixed monotone if there exists a *decomposition function* that separates the vector field of the system into solely increasing and solely decreasing components [1], [5]–[8]. Said decomposition function is used to construct a $2n$–dimensional *embedding system* that is monotone with respect to a particular southeast order (see Section II). This allows for the application of tools from monotone systems theory to the embedding system dynamics, which yields conclusions on the reachability and safety properties of the original system. In particular, if the original system is subject to a $p$–dimensional disturbance input, the resulting embedding system is subject to a $2p$–dimensional disturbance input; considering the worst-case inputs of these disturbances allows for the efficient computation of reachable and forward invariant sets of the original system [4].

Prior works on mixed monotone systems did not consider disturbances arising from unknown but state-dependent uncertainty in the dynamics, although this is a common practical scenario. To that end, GPs have been used to model unknown functions to great effect in statistics and machine learning [9], as they are able to model distributions over any continuous domain and provide confidence estimates over a given range of function values, even with few observations. One can consequently take advantage of these confidence estimates to form high-confidence bounds on the disturbance and update these bounds as more observations are gathered.

The advantages afforded by GPs, and learning methods in general, in estimating unknown functions have not gone unnoticed by the controls community; [10] provides a method for online tuning of controller parameters using GPs while fulfilling safety criteria, [11] describes a method for incorporating Reinforcement Learning using GPs into classical model reference adaptive control, [12], [13] explore the coverage control problem for estimating unknown spatial fields using GPs, and [14] derives a uniform error bound for GPs that is used to calculate safety bounds for unknown dynamical systems. In service of providing high probability safety guarantees, [15] uses Bayesian learning to obtain a distribution over the system dynamics, [16] presents a model predictive control formulation that incorporates GPs, and [17] implements reinforcement learning to model uncertainties within control barrier function and control Lyapunov function constraints. Further, a min-norm control Lyapunov function-based stabilizing controller for control affine systems with uncertain dynamics utilizing GP regression is presented in [18]. Additionally, much work has been done in leveraging Hamilton-Jacobi reachability methods for safety. The paper [19] provides a least-restrictive safety-preserving control framework based on combining these methods with GPs, and [20] synthesizes techniques to speed up computation of Hamilton-Jacobi safe sets as uncertainties are reduced.

We present here a method for computing reachable and forward invariant sets for systems whose dynamics include unknown components, drawing from the previous literature on mixed monotone systems. We accomplish this by assuming bounds on the unknown components of the system and then show that these bounds lead to the identification of reachable and forward invariant sets, resulting in computationally efficient algorithms. We then show that, when GPs are used to model the unknown components, this assumption is particularly appropriate, as it enables the calculation of bounds that are correct with high probability.

The proposed approach is well suited to high dimensional systems with low dimensional uncertainty that appears as unknown components in the dynamics. This scenario often occurs in practice, for example, in mechanical systems, where uncertain forces generally affect only the velocity dynamics and might be a function of only positional state. For such systems, bounds on the low dimensional uncertainty can be efficiently evaluated using update laws for GPs and direct sampling, while the theory of mixed monotone systems accommodates the higher dimensional dynamics, leading to tractable computations. In addition, in contrast to much of the prior work using GPs in control systems, we do not assume the dynamics are affine in the unknown components, and we model systems in continuous time.

Theorems 1, 3, 4, and 7 have previously appeared in conference proceedings [21], which focused on demonstrating the validity of the proposed formulation in calculating high probability reachable and forward invariant sets. We expand upon the capabilities of this formulation with additional theoretical results proving that calculated sets tighten as the bounds on the unknown component tighten, as well as a method for reducing the space over which the GPs need to be sampled to the border of the hyperrectangular sets. These additional results are showcased in a new set of case studies.

This paper is organized as follows. Key notation is introduced in Section II. In Section III, we formally define the assumptions made and the problems to be solved. Subsequently, in Section IV, we illustrate the key theoretical results that solve the previously defined problems, before detailing the theory that allows us to leverage GPs in Section V. In Section VI, we showcase several demonstrations of the results, and we conclude the paper in Section VII.

## II. Notation

Let $(x, y)$ denote the vector concatenation of $x, y \in \mathbb{R}^n$, i.e., $(x, y) := [x^T \ y^T]^T \in \mathbb{R}^{2n}$. Additionally, $\preceq$ denotes the componentwise vector order, i.e., $x \preceq y$ if and only if $x_i \leq y_i$ for all $i \in \{1, ..., n\}$ where vector components are indexed via subscript.

Given $x, y \in \mathbb{R}^n$ such that $x \preceq y$, we denote the hyperrectangle defined by the endpoints $x$ and $y$ using the notation $[x, y] := \{z \in \mathbb{R}^n \mid x \preceq z \text{ and } z \preceq y\}$. Also, given $a = (x, y) \in \mathbb{R}^{2n}$ with $x \preceq y$, $[\![a]\!]$ denotes the hyperrectangle formed by the first and last $n$ components of $a$, i.e., $[\![a]\!] := [x, y]$.

Finally, let $\preceq_{\text{SE}}$ denote the *southeast order* on $\mathbb{R}^{2n}$ defined by $(x, x') \preceq_{\text{SE}} (y, y')$ if and only if $x \preceq y$ and $y' \preceq x'$. In particular, observe that when $x \preceq x'$ and $y \preceq y'$,

$$(x, x') \preceq_{\text{SE}} (y, y') \iff [y, y'] \subseteq [x, x']. \tag{1}$$

## III. Problem Formulation

Consider the continuous-time, Lipschitz dynamical system

$$\dot{x} = f(x, w) \tag{2}$$

where $x \in \mathbb{R}^n$ is the system state and $w \in \mathbb{R}^p$ is an unknown, state-dependent component of the dynamics so that $w_i = g_i(x)$ where $g_i$ is unknown. For example, $g_i$ might account for higher order nonlinearities not explicitly captured in the model.

We make the following two fundamental assumptions throughout.

**Assumption 1.** *Each* $w_i$*,* $i \in \{1,\ldots,p\}$ *in* (2) *is state-dependent so that* $w_i = g_i(x)$ *for some unknown, Lipschitz continuous* $g_i : \mathbb{R}^n \mapsto \mathbb{R}$*.*

*Further, there exist known, Lipschitz continuous functions* $\underline{\gamma}_i(x, \widehat{x})$ *and* $\overline{\gamma}_i(x, \widehat{x})$*,* $\underline{\gamma}_i, \overline{\gamma}_i : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$*, for all* $i \in \{1,\ldots,p\}$ *such that*

$$\underline{\gamma}_i(\underline{x}, \overline{x}) \leq g_i(x) \leq \overline{\gamma}_i(\underline{x}, \overline{x}) \; \forall x \in [\underline{x}, \overline{x}] \tag{3}$$

*for all* $\underline{x}, \overline{x} \in \mathbb{R}^n$ *with* $\underline{x} \preceq \overline{x}$*. Without loss of generality, we assume* $\underline{\gamma}_i$ *and* $\overline{\gamma}_i$ *satisfy the natural inclusion property that for all* $\underline{x} \preceq \underline{y} \preceq \overline{y} \preceq \overline{x}$ *it holds that* $\underline{\gamma}_i(\underline{x}, \overline{x}) \leq \underline{\gamma}_i(\underline{y}, \overline{y}) \leq \overline{\gamma}_i(\underline{y}, \overline{y}) \leq \overline{\gamma}_i(\underline{x}, \overline{x})$*.*

In Section V, we present a method for computing functions $\underline{\gamma}_i(x, \widehat{x})$ and $\overline{\gamma}_i(x, \widehat{x})$ satisfying Assumption 1 with high probability when $g_i$ is modeled as a GP, and we explicitly quantify the probability that (3) holds.

We denote by $g(x)$, $\underline{\gamma}$, and $\overline{\gamma}$ the vector concatenation of $g_i$, $\underline{\gamma}_i$, and $\overline{\gamma}_i$ for $i = 1,\ldots,p$.

**Assumption 2.** *The system* (2) *is mixed monotone with respect to a Lipschitz decomposition function* $\delta(x, w, \widehat{x}, \widehat{w})$*, that is,* $\delta : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R}^p \mapsto \mathbb{R}^n$ *satisfies:*

1) *For all* $x$ *and all* $w$*,* $\delta(x, w, x, w) = f(x, w)$*;*
2) *For all* $i$*,* $j \in \{1, \cdots, n\}$*,* $i \neq j$*,* $\frac{\partial \delta_i}{\partial x_j}(x, w, \widehat{x}, \widehat{w}) \geq 0$ *for all* $x$*,* $\widehat{x}$ *and all* $w, \widehat{w}$ *wherever the derivative exists;*
3) *For all* $i$*,* $j \in \{1, \cdots, n\}$*,* $\frac{\partial \delta_i}{\partial \widehat{x}_j}(x, w, \widehat{x}, \widehat{w}) \leq 0$ *for all* $x, \widehat{x}$ *and all* $w, \widehat{w}$ *wherever the derivative exists; and,*
4) *For all* $i \in \{1, \cdots, n\}$ *and all* $k \in \{1, \cdots, p\}$*,* $\frac{\partial \delta_i}{\partial w_k}(x, w, \widehat{x}, \widehat{w}) \geq 0$ *and* $\frac{\partial \delta_i}{\partial \widehat{w}_k}(x, w, \widehat{x}, \widehat{w}) \leq 0$ *for all* $x, \widehat{x}$ *and all* $w, \widehat{w}$ *wherever the derivative exists.*

Assumption 2 is not particularly restrictive as it has been shown in [4] that all systems with Lipschitz dynamics are mixed monotone with respect to some decomposition function. Generally, domain-specific knowledge is leveraged to obtain closed-form decomposition functions, and for some classes of systems such as those for which the partial derivatives of $f$ are bounded, decomposition functions are readily constructed from $f$. Additionally, the choice of decomposition function affects the conservatism of the reachable set overapproximation [22]. For examples of decomposition functions, see equation (79) in Section VI, or [1] and citations therein for more details. Lastly, we note that for simplicity, we take $\mathbb{R}^n$ as the domain, although it is possible to restrict to a domain that is a hyperrectangular subset of $\mathbb{R}^n$, as described in [1].

For initial condition $x_0 \in \mathbb{R}^n$, let $\phi(t, x_0)$ denote the resulting state trajectory of (2) when $w = g(x)$, that is, $\phi(t, x_0)$ satisfies $\frac{d}{dt}\phi(t, x_0) = f(\phi(t, x_0), g(\phi(t, x_0)))$. Given some $X_0 \subseteq \mathbb{R}^n$, the *$T$-horizon reachable set from $X_0$* for (2) is the set of states reachable over the time horizon $T$ from any initial condition $x_0 \in X_0$ and is denoted

$$R(T, X_0) = \{\phi(T, x_0) \mid x_0 \in X_0\}. \tag{4}$$

Even when the dynamics are fully known, computing exact reachable sets is generally not possible. Thus, we are interested in computing approximations of reachable sets. In particular, over-approximations are often preferred for, e.g., safety verification. A key feature of mixed monotone systems is that hyperrectangular over-approximations of reachable sets are efficiently computed from trajectories of a $2n$ dimensional embedding system constructed from the decomposition function $\delta$. In Section IV, we extend this fundamental property to systems with state-dependent uncertainty satisfying Assumption 1, expanding on our results from [21]. The focus of this paper is on the tractable computation of over-approximations for reachable sets of (2).

**Problem 1.** *Given* $X_0 = [\underline{x}_0, \overline{x}_0]$ *for some* $\underline{x}_0, \overline{x}_0 \in \mathbb{R}^n$ *with* $\underline{x}_0 \preceq \overline{x}_0$*, compute* $\widehat{R}(T, X_0)$ *such that* $R(T, X_0) \subseteq \widehat{R}(T, X_0)$*.*

Closely related to the problem of computing reachable sets is the problem of computing forward invariant sets for (2).

**Problem 2.** *Identify sets* $A \subseteq \mathbb{R}^n$ *such that* $R(t, A) \subseteq A$ *for all* $t \geq 0$*.*

In Section IV, we present solutions to Problems 1 and 2 using mixed monotone systems theory. Then, in Section V, we use the theory of GPs to derive a general approach to satisfying Assumption 1 and, in particular, obtaining the necessary functions $\underline{\gamma}$ and $\overline{\gamma}$ such that (3), and consequently the identified reachable and forward invariant sets, holds with high probability.

## IV. Reachable and Invariant Sets

We present solutions to Problems 1 and 2, including results that show improvements to the tightness of the calculated sets as bounds on the unknown components are tightened. We then outline an additional refinement to our formulation that results in more accurate reachable and forward invariant sets at the cost of increased computational complexity.

We begin by recalling the fundamental result of mixed monotone systems theory (see, *e.g.*, [1]). Construct the *embedding system* with state $(x, \widehat{x}) \in \mathbb{R}^n \times \mathbb{R}^n$ and disturbance $(w, \widehat{w}) \in \mathbb{R}^p \times \mathbb{R}^p$

$$\begin{bmatrix} \dot{x} \\ \dot{\widehat{x}} \end{bmatrix} = \varepsilon(x, w, \widehat{x}, \widehat{w}) := \begin{bmatrix} \delta(x, w, \widehat{x}, \widehat{w}) \\ \delta(\widehat{x}, \widehat{w}, x, w) \end{bmatrix}. \tag{5}$$

Denote the state of (5) at time $t$ when initialized at $(\underline{x}_0, \overline{x}_0)$ under some disturbance input signal $(w, \widehat{w}) : [0, \infty) \to \mathbb{R}^p \times \mathbb{R}^p$ by $\Phi^\varepsilon(t; (\underline{x}_0, \overline{x}_0), (w, \widehat{w}))$. The fundamental result of mixed monotone systems theory is that (5) is a monotone control system as defined in [23] with respect to the southeast order on state and disturbance; that is, given $a, a' \in \mathbb{R}^n \times \mathbb{R}^n$ and $b, b' : [0, \infty) \to \mathbb{R}^p \times \mathbb{R}^p$ such that $a \preceq_{\mathrm{SE}} a'$ and $b(t) \preceq_{\mathrm{SE}} b'(t)$ for all $t \geq 0$, then for all $t \geq 0$,

$$\Phi^\varepsilon(t; a, b) \preceq_{\mathrm{SE}} \Phi^\varepsilon(t; a', b'). \tag{6}$$

### A. Reachable Sets

Our main result for calculating reachable sets is as follows.

**Theorem 1.** *Consider* (2) *satisfying Assumptions 1 and 2. Let* $X_0 = [\underline{x}_0, \overline{x}_0]$ *for* $\underline{x}_0, \overline{x}_0 \in \mathbb{R}^n$*, with* $\underline{x}_0 \preceq \overline{x}_0$*, be a* *hyperrectangular set of initial conditions. Let* $(x(t), \widehat{x}(t))$ *be the solution to the* $2n$ *dimensional system*

$$\begin{bmatrix}\dot{x}\\ \dot{\widehat{x}}\end{bmatrix} = e(x,\widehat{x}) := \begin{bmatrix}\delta(x,\underline{\gamma}(x,\widehat{x}),\widehat{x},\overline{\gamma}(x,\widehat{x}))\\ \delta(\widehat{x},\overline{\gamma}(x,\widehat{x}),x,\underline{\gamma}(x,\widehat{x}))\end{bmatrix} \tag{7}$$

*with initial condition* $(x(0), \widehat{x}(0)) = (\underline{x}_0, \overline{x}_0)$. *Then,* $R(T, X_0) \subseteq [x(T), \widehat{x}(T)]$ *for all* $T \geq 0$.

*Proof:*
We first construct the embedding system (5), then substitute $(\underline{\gamma}(x,\widehat{x}), \overline{\gamma}(x,\widehat{x}))$ for $(w, \widehat{w})$, giving the embedding system (7). We denote the solutions $(x(t), \widehat{x}(t))$ of this system by $\Phi^e(t; (\underline{x}_0, \overline{x}_0))$, which we note is equivalent to $\Phi^\varepsilon(t; (\underline{x}_0, \overline{x}_0), (\underline{\gamma}, \overline{\gamma}))$, where it is understood that $(\underline{\gamma}, \overline{\gamma})$ is evaluated along the $(x, \widehat{x})$ trajectory. Note also that

$$\Phi^\varepsilon(t; (x_0, x_0), (g, g)) = (\phi(t, x_0), \phi(t, x_0)), \tag{8}$$

where it is understood that $(g, g)$ is evaluated along the system trajectory, *i.e.*, $\Phi^\varepsilon(t; (x_0, x_0), (g, g))$ is equivalent to two copies of state trajectories of the original dynamics (2).

Now, Assumption 1 equivalently states that

$$(\underline{\gamma}(x,\widehat{x}), \overline{\gamma}(x,\widehat{x})) \preceq_{\text{SE}} (g(x'), g(x'))\ \forall x \preceq x' \preceq \widehat{x}. \tag{9}$$

Moreover, for any initial state $x_0 \in [\underline{x}_0, \overline{x}_0]$, we have equivalently, $(\underline{x}_0, \overline{x}_0) \preceq_{\text{SE}} (x_0, x_0)$. Per the fundamental result of mixed monotone systems theory (6), we have that (9) implies

$$\Phi^e(t; (\underline{x}_0, \overline{x}_0)) \preceq_{\text{SE}} \Phi^\varepsilon(t; (x_0, x_0), (g, g)) \tag{10}$$

for all $t \geq 0$. By (8),

$$\phi(t, x_0) \in [\![\Phi^e(t; (\underline{x}_0, \overline{x}_0))]\!] \tag{11}$$

for all $t \geq 0$. Finally, recalling (4) gives us

$$R(T, X_0) \subseteq [\![\Phi^e(T; (\underline{x}_0, \overline{x}_0))]\!] = [x(T), \widehat{x}(T)]. \tag{12}$$

■

Theorem 1 establishes that $\widehat{R}(T, X_0) := [x(T), \widehat{x}(T)]$, where $x(T), \widehat{x}(T)$ are obtained as the solution to the ODE (7), provides a hyperrectangular over-approximation of the true reachable set $R(T, X_0)$, thus solving Problem 1.

As detailed in the next section, the theory of GPs naturally leads to a methodology for modeling and obtaining $\underline{\gamma}$ and $\overline{\gamma}$ satisfying Assumption 2. In this case, it is further reasonable to envision updating the bounds $\underline{\gamma}$ and $\overline{\gamma}$ by incorporating newly collected data into estimation of the GPs. In this case, estimates of reachable sets become tighter, provided that the updated bounds $\underline{\gamma}$ and $\overline{\gamma}$ become tighter. This is formalized in Theorem 2.

**Theorem 2.** *Consider* (2) *satisfying Assumption 2. Let* $\underline{\gamma}, \overline{\gamma}$ *be a pair of bounds satisfying Assumption 1, and let* $\underline{\gamma}', \overline{\gamma}'$ *be another pair of bounds satisfying Assumption 1. Suppose further that the bounds* $\underline{\gamma}', \overline{\gamma}'$ *are tighter than* $\underline{\gamma}, \overline{\gamma}$, *that is, for all* $\underline{x}, \overline{x}$ *with* $\underline{x} \preceq \overline{x}$, *it holds that*

$$\underline{\gamma}(\underline{x}, \overline{x}) \preceq \underline{\gamma}'(\underline{x}, \overline{x}) \preceq \overline{\gamma}'(\underline{x}, \overline{x}) \preceq \overline{\gamma}(\underline{x}, \overline{x}). \tag{13}$$

*Construct the embedding systems*

$$\begin{bmatrix}\dot{x}\\ \dot{\widehat{x}}\end{bmatrix} = e'(x,\widehat{x}) := \begin{bmatrix}\delta(x,\underline{\gamma}'(x,\widehat{x}),\widehat{x},\overline{\gamma}'(x,\widehat{x}))\\ \delta(\widehat{x},\overline{\gamma}'(x,\widehat{x}),x,\underline{\gamma}'(x,\widehat{x}))\end{bmatrix} \tag{14}$$

*and*

$$\begin{bmatrix}\dot{x}\\ \dot{\widehat{x}}\end{bmatrix} = e(x,\widehat{x}) := \begin{bmatrix}\delta(x,\underline{\gamma}(x,\widehat{x}),\widehat{x},\overline{\gamma}(x,\widehat{x}))\\ \delta(\widehat{x},\overline{\gamma}(x,\widehat{x}),x,\underline{\gamma}(x,\widehat{x}))\end{bmatrix}. \tag{15}$$

*Then*

$$[\![\Phi^{e'}(T; (\underline{x}_0, \overline{x}_0)]\!] \subseteq [\![\Phi^e(T; (\underline{x}_0, \overline{x}_0)]\!]. \tag{16}$$

*Proof:*
We first note that (13) equivalently states

$$(\underline{\gamma}(\underline{x}, \overline{x}), \overline{\gamma}(\underline{x}, \overline{x})) \preceq_{\text{SE}} (\underline{\gamma}'(\underline{x}, \overline{x}), \overline{\gamma}'(\underline{x}, \overline{x}))\ \forall \underline{x} \preceq \overline{x}. \tag{17}$$

It thus follows that, per the fundamental result of mixed monotone systems theory (6),

$$\begin{aligned}&\Phi^\varepsilon(t; (x,\widehat{x}), (\underline{\gamma}(x,\widehat{x}), \overline{\gamma}(x,\widehat{x})))\\ &\preceq_{\text{SE}}\ \Phi^\varepsilon(t; (x,\widehat{x}), (\underline{\gamma}'(x,\widehat{x}), \overline{\gamma}'(x,\widehat{x}))),\end{aligned} \tag{18}$$

or, equivalently,

$$\Phi^e(T; (\underline{x}_0, \overline{x}_0)) \preceq_{\text{SE}} \Phi^{e'}(T; (\underline{x}_0, \overline{x}_0)), \tag{19}$$

which is equivalent to (16). ■

The first case study of Section VI uses Theorem 2 to learn a tight overapproximation of dynamics subject to a sudden unknown disturbance by collecting observations of the unknown component's behavior.

### *B. Forward Invariant Sets*

The same mixed monotone formulation allows for the efficient computation of forward invariant sets.

**Theorem 3.** *Consider* (2) *satisfying Assumptions 1 and 2. If* $\exists\ \underline{x}^*, \overline{x}^*, \underline{w}^*, \overline{w}^*$ *with* $\underline{x}^* \preceq \overline{x}^*$ *and* $\underline{w}^* \preceq \overline{w}^*$ *such that*

$$\underline{w}^* \preceq \underline{\gamma}(\underline{x}^*, \overline{x}^*) \text{ and } \overline{\gamma}(\underline{x}^*, \overline{x}^*) \preceq \overline{w}^* \tag{20}$$

*and*

$$\delta(\underline{x}^*, \underline{w}^*, \overline{x}^*, \overline{w}^*) \succeq 0 \text{ and } \delta(\overline{x}^*, \overline{w}^*, \underline{x}^*, \underline{w}^*) \preceq 0, \tag{21}$$

*then* $[\underline{x}^*, \overline{x}^*]$ *is forward invariant for* (2).

*Proof:*
We construct an embedding system similar to the system outlined by (7), substituting $(\underline{w}^*, \overline{w}^*)$ for $(\underline{\gamma}(x,\widehat{x}), \overline{\gamma}(x,\widehat{x}))$:

$$\begin{bmatrix}\dot{x}\\ \dot{\widehat{x}}\end{bmatrix} = e^*(x,\widehat{x}) := \begin{bmatrix}\delta(x,\underline{w}^*,\widehat{x},\overline{w}^*)\\ \delta(\widehat{x},\overline{w}^*,x,\underline{w}^*)\end{bmatrix}. \tag{22}$$

The *upper triangle* of this embedding system is defined as $\mathcal{T} := \{(x,\widehat{x}) \in \mathbb{R}^n \times \mathbb{R}^n |\ x \preceq \widehat{x}\}$, and the set of points in $\mathcal{T}$ such that the vector field of the embedding system points into the southeast cone is defined as $\mathcal{S} := \{(x,\widehat{x}) \in \mathcal{T} |\ 0 \preceq_{\text{SE}} e^*(x,\widehat{x})\}$. The set $\mathcal{T}$ is called the upper triangle as it can be visualized as the set of states above the $x = \widehat{x}$ line in the sense of the partial order $\leq$. Per [4, Lemma 1], $\mathcal{T}$ is forward invariant for (22). Further, as a direct result of [24, Ch. 3, Prop 2.1], $\mathcal{S}$ is also forward invariant for (22) and $\Phi^{e^*}(t_1; a) \preceq_{\text{SE}} \Phi^{e^*}(t_2; a)$ for all $a \in \mathcal{S}$ and all $0 \leq t_1 \leq t_2$.

In particular, $a \preceq_{\rm SE} \Phi^{e^*}(t;a)$ for all $a \in \mathcal{S}$ and all $t \geq 0$, i.e., by (1), $[\![\Phi^{e^*}(t;a)]\!] \subseteq [\![a]\!]$ for all $a \in \mathcal{S}$ and all $t \geq 0$.

Let $a^* := (\underline{x}^*, \overline{x}^*)$. By (21), $a^* \in \mathcal{S}$, so that $[\![\Phi^{e^*}(t;a^*)]\!] \subseteq [\![a^*]\!]$ for all $t \geq 0$. From Theorem 1, for any $t \geq 0$, we have that $R(t, [\![a^*]\!]) \subseteq [\![\Phi^e(t;a^*)]\!]$ where $\Phi^e$ is as defined in the proof of Theorem 1. Next, (20) and Theorem 2 implies $[\![\Phi^e(t;a^*)]\!] \subseteq [\![\Phi^{e^*}(t;a^*)]\!]$. Combined, we have $R(t, [\![a^*]\!]) \subseteq [\![\Phi^e(t;a^*)]\!] \subseteq [\![\Phi^{e^*}(t;a^*)]\!] \subseteq [\![a^*]\!]$ for all $t \geq 0$, i.e., $[\![a^*]\!] = [\underline{x}^*, \overline{x}^*]$ is forward invariant for (2). ■

We immediately have the following corollary.

**Corollary 1.** *Consider* (2) *satisfying Assumptions 1 and 2. If* $(\underline{x}^*, \overline{x}^*)$ *is an equilibrium for the embedding system* (7)*, then* $[\underline{x}^*, \overline{x}^*]$ *is forward invariant for* (2)*.*

Corollary 1 suggests a numerical method for identifying invariant sets for (2): compute equilibria of the embedding system (7) by, *e.g.*, initializing the embedding system dynamics at some point and simulating the dynamics to determine if the trajectory converges. One difficulty of this method, however, is that the functions $\underline{\gamma}$ and $\overline{\gamma}$ must be evaluated at each point along the entire trajectory. As discussed in the next section, this is computationally reasonable in some cases, but may be impractical in other cases.

Instead, Theorem 3 offers an alternative method for identifying invariant sets that relies on evaluating $\underline{\gamma}$ and $\overline{\gamma}$ at a sequence of points and simulating the resulting embedding dynamics with $\underline{\gamma}$ and $\overline{\gamma}$ fixed to these evaluations.

To that end, consider an initial $\underline{x}^1 \preceq \overline{x}^1$ and construct the sequences $\{\underline{x}^k\}_{k=1}^\infty$, $\{\overline{x}^k\}_{k=1}^\infty$ according to the recursion

$$\underline{w}^k := \underline{\gamma}(\underline{x}^k, \overline{x}^k),\ \overline{w}^k := \overline{\gamma}(\underline{x}^k, \overline{x}^k) \tag{23}$$

$$(\underline{x}^{k+1}, \overline{x}^{k+1}) := \lim_{t\to\infty} (x^k(t), \widehat{x}^k(t)) \tag{24}$$

provided the limit exists, where $(x^k(t), \widehat{x}^k(t))$ solves

$$\dot{x}^k = \delta(x^k, \underline{w}^k, \widehat{x}^k, \overline{w}^k),\ \dot{\widehat{x}}^k = \delta(\widehat{x}^k, \overline{w}^k, x^k, \underline{w}^k) \tag{25}$$

with initial condition $x^k(0) = \underline{x}^k$, $\widehat{x}^k(0) = \overline{x}^k$.

**Theorem 4.** *Consider* (2) *satisfying Assumptions 1 and 2. For any initial* $\underline{x}^1 \preceq \overline{x}^1$*, if the sequences* $\{\underline{x}^k\}_{k=1}^\infty$, $\{\overline{x}^k\}_{k=1}^\infty$ *constructed according to* (23)–(25) *are well-defined (i.e. the limits exist) and converge, then* $(\underline{x}^*, \overline{x}^*) := \lim_{k\to\infty}(\underline{x}^k, \overline{x}^k)$ *constitutes an equilibrium of* (7) *and, hence,* $[\underline{x}^*, \overline{x}^*]$ *is forward invariant for* (2)*.*

*In addition, if for some* $K$ *it holds that*

$$\underline{x}^K \preceq \underline{x}^{K+1} \text{ and } \overline{x}^{K+1} \preceq \overline{x}^K, \tag{26}$$

*then for all* $k \geq K$*, it holds that* $\underline{x}^k \preceq \underline{x}^{k+1}$ *and* $\overline{x}^{k+1} \preceq \overline{x}^k$ *and the sequences* $\{\underline{x}^k\}_{k=1}^\infty$, $\{\overline{x}^k\}_{k=1}^\infty$ *converge. Moreover, each* $[\underline{x}^k, \overline{x}^k]$ *for* $k \geq K$ *is forward invariant for* (2)*.*

The proof of Theorem 4 relies in part on the following lemma, which is an immediate consequence of Property 4 of the decomposition function $\delta$ stated in Assumption 2.

**Lemma 1.** *Consider* (2) *satisfying Assumptions 1 and 2. Suppose* $\underline{x}^*, \overline{x}^*, \underline{w}^*, \overline{w}^*$ *with* $\underline{x}^* \preceq \overline{x}^*$ *and* $\underline{w}^* \preceq \overline{w}^*$ *are such that*

$$\delta(\underline{x}^*, \underline{w}^*, \overline{x}^*, \overline{w}^*) = 0 \text{ and } \delta(\overline{x}^*, \overline{w}^*, \underline{x}^*, \underline{w}^*) = 0. \tag{27}$$

*If* $\underline{w}'$ *and* $\overline{w}'$ *satisfy* $\underline{w}^* \preceq \underline{w}' \preceq \overline{w}' \preceq \overline{w}^*$*, then*

$$\delta(\underline{x}^*, \underline{w}', \overline{x}^*, \overline{w}') \succeq 0 \text{ and } \delta(\overline{x}^*, \overline{w}', \underline{x}^*, \underline{w}') \preceq 0. \tag{28}$$

*Proof of Theorem 4:*
For the first part of Theorem 4, construct the embedding system described by (25) for each iteration $k$,

$$\begin{bmatrix} \dot{x}^k \\ \dot{\widehat{x}}^k \end{bmatrix} = e^k(x^k, \widehat{x}^k) := \begin{bmatrix} \delta(x^k, \underline{w}^k, \widehat{x}^k, \overline{w}^k) \\ \delta(\widehat{x}^k, \overline{w}^k, x^k, \underline{w}^k) \end{bmatrix}. \tag{29}$$

Since the limit $\lim_{t\to\infty}(x^k(t), \widehat{x}^k(t))$ exists, $(\underline{x}^{k+1}, \overline{x}^{k+1})$ as defined by (24) must be such that

$$\begin{bmatrix} \dot{x} \\ \dot{\widehat{x}} \end{bmatrix} = e^k(\underline{x}^{k+1}, \overline{x}^{k+1}) = 0, \tag{30}$$

i.e., $(\underline{x}^{k+1}, \overline{x}^{k+1})$ is an equilibrium for $e^k$.

Because the sequences $\{\underline{x}^k\}_{k=1}^\infty$, $\{\overline{x}^k\}_{k=1}^\infty$ converge, and because $\delta$ is Lipschitz continuous by Assumption 2, the Lipschitz continuous function $e^*(x, \widehat{x}) = \begin{bmatrix}\delta(x, \underline{w}^*, \widehat{x}, \overline{w}^*) & \delta(\widehat{x}, \overline{w}^*, x, \underline{w}^*)\end{bmatrix}^T$ where $(\underline{w}^*, \overline{w}^*) := \lim_{k\to\infty}(\underline{w}^k, \overline{w}^k)$ satisfies $e^*(\underline{x}^*, \overline{x}^*) = 0$ where $(\underline{x}^*, \overline{x}^*) := \lim_{k\to\infty}(\underline{x}^k, \overline{x}^k)$. In turn, this means that $(\underline{x}^*, \overline{x}^*)$ is an equilibrium of the embedding system $\begin{bmatrix}\dot{x} & \dot{\widehat{x}}\end{bmatrix}^T = e^*(x, \widehat{x})$, which we recall from Corollary 1 means that $[\underline{x}^*, \overline{x}^*]$ is forward invariant for (2).

For the second part of Theorem 4, we note from the inclusion property in Assumption 2 that, for the $K$ at which (26) holds, it must be true that $\underline{\gamma}(\underline{x}^K, \overline{x}^K) \preceq \underline{\gamma}(\underline{x}^{K+1}, \overline{x}^{K+1}) \preceq \overline{\gamma}(\underline{x}^{K+1}, \overline{x}^{K+1}) \preceq \overline{\gamma}(\underline{x}^K, \overline{x}^K)$, or, equivalently, $\underline{w}^K \preceq \underline{w}^{K+1} \preceq \overline{w}^{K+1} \preceq \overline{w}^K$. From (30), it also holds that $\delta(\underline{x}^{K+1}, \underline{w}^K, \overline{x}^{K+1}, \overline{w}^K) = 0$ and $\delta(\overline{x}^{K+1}, \overline{w}^K, \underline{x}^{K+1}, \underline{w}^K) = 0$. Thus, by Lemma 1,

$$\delta(\underline{x}^{K+1}, \underline{w}^{K+1}, \overline{x}^{K+1}, \overline{w}^{K+1}) \succeq 0, \tag{31}$$

$$\delta(\overline{x}^{K+1}, \overline{w}^{K+1}, \underline{x}^{K+1}, \underline{w}^{K+1}) \preceq 0, \tag{32}$$

which subsequently means that $\underline{x}^{K+1} \preceq \underline{x}^{K+2}$ and $\overline{x}^{K+2} \preceq \overline{x}^{K+1}$ must also hold, which means that $\underline{w}^{K+1} \preceq \underline{w}^{K+2} \preceq \overline{w}^{K+2} \preceq \overline{w}^{K+1}$ also holds, and so on by induction. This also extends back one step to iteration $K$, i.e., it must also have been true that $\delta(\underline{x}^K, \underline{w}^K, \overline{x}^K, \overline{w}^K) \succeq 0$ and $\delta(\overline{x}^K, \overline{w}^K, \underline{x}^K, \underline{w}^K) \preceq 0$ in order for (26) to hold in the first place (though we note that $\underline{w}^{K-1} \preceq \underline{w}^K \preceq \overline{w}^K \preceq \overline{w}^{K-1}$ does not necessarily have to hold for this to be the case).

As a result, for all $k \geq K$, it holds that $\underline{x}^k \preceq \underline{x}^{k+1}$ and $\overline{x}^{k+1} \preceq \overline{x}^k$ and the sequences $\{\underline{x}^k\}_{k=1}^\infty$, $\{\overline{x}^k\}_{k=1}^\infty$ converge. Additionally, as $\delta(\underline{x}^k, \underline{w}^k, \overline{x}^k, \overline{w}^k) \succeq 0$ and $\delta(\overline{x}^k, \overline{w}^k, \underline{x}^k, \underline{w}^k) \preceq 0$ hold for all $k \geq K$, by Theorem 3 each $[\underline{x}^k, \overline{x}^k]$ for $k \geq K$ is forward invariant for (2). ■

Theorem 3 and Lemma 1 lead to another consequence regarding identifying forward invariant sets. As previously established, it is reasonable to envision updating the bounds $\underline{\gamma}$ and $\overline{\gamma}$ by incorporating newly collected data into estimation of the GPs. As a result, estimates of any forward invariant sets can be used to establish new, tighter forward invariant sets, provided that the updated bounds $\underline{\gamma}$ and $\overline{\gamma}$ become tighter. This is formalized in Proposition 1 below.

**Proposition 1.** *Consider* (2) *satisfying Assumption 2. Let* $\underline{\gamma}, \overline{\gamma}$ *be one pair of bounds satisfying Assumption 1, and let* $\underline{\gamma}', \overline{\gamma}'$ *be another pair of bounds satisfying Assumption 1. Suppose further that the bounds* $\underline{\gamma}, \overline{\gamma}$ *are tighter than* $\underline{\gamma}', \overline{\gamma}'$*, that is, for all* $\underline{x}, \overline{x}$ *with* $\underline{x} \preceq \overline{x}$*, it holds that*

$$\underline{\gamma}'(\underline{x}, \overline{x}) \preceq \underline{\gamma}(\underline{x}, \overline{x}) \preceq \overline{\gamma}(\underline{x}, \overline{x}) \preceq \overline{\gamma}'(\underline{x}, \overline{x}). \tag{33}$$

*Let* $\underline{x}^{*\prime}, \overline{x}^{*\prime}$ *satisfy*

$$0 = \delta(\underline{x}^{*\prime}, \underline{\gamma}'(\underline{x}^{*\prime}, \overline{x}^{*\prime}), \overline{x}^{*\prime}, \overline{\gamma}'(\underline{x}^{*\prime}, \overline{x}^{*\prime})) \tag{34}$$

$$0 = \delta(\overline{x}^{*\prime}, \overline{\gamma}'(\underline{x}^{*\prime}, \overline{x}^{*\prime}), \underline{x}^{*\prime}, \underline{\gamma}'(\underline{x}^{*\prime}, \overline{x}^{*\prime})), \tag{35}$$

*that is,* $\underline{x}^{*\prime}, \overline{x}^{*\prime}$ *is an equilibrium for the embedding dynamics constructed using* $\underline{\gamma}'$ *and* $\overline{\gamma}'$*. Construct the sequences* $\{\underline{x}^k\}_{k=1}^{\infty}$, $\{\overline{x}^k\}_{k=1}^{\infty}$ *as prescribed in* (23)–(25) *with initial* $\underline{x}^1 = \underline{x}^{*\prime}$ *and* $\overline{x}^1 = \overline{x}^{*\prime}$*. Then*

$$\underline{x}^k \preceq \underline{x}^{k+1} \text{ and } \overline{x}^{k+1} \preceq \overline{x}^k \tag{36}$$

*for all* $k \geq 1$ *so that, in particular, the conclusions of Theorem 4 hold for all* $k \geq 1$*, namely, the sequences* $\{\underline{x}^k\}_{k=1}^{\infty}$, $\{\overline{x}^k\}_{k=1}^{\infty}$ *converge and each* $[\underline{x}^k, \overline{x}^k]$ *is forward invariant for* (2)*.*

*Proof:*
Given (33), $(\underline{x}^*, \overline{x}^*) = (\underline{x}^{*\prime}, \overline{x}^{*\prime})$, $(\underline{w}^*, \overline{w}^*) = (\underline{\gamma}', \overline{\gamma}')$, and $(\underline{w}^{*\prime}, \overline{w}^{*\prime}) = (\underline{\gamma}, \overline{\gamma})$ satisfy the conditions for Lemma 1. Thus, from Theorem 4 it holds that

$$\delta(\underline{x}^1, \underline{w}^1, \overline{x}^1, \overline{w}^1) \succeq 0,\ \delta(\overline{x}^1, \overline{w}^1, \underline{x}^1, \underline{w}^1) \preceq 0. \tag{37}$$

The rest of the proof follows from the second part of the proof of Theorem 4. ∎

Theorem 3 and Corollary 1 indicate that the hyperrectangular set $A := [\underline{x}^*, \overline{x}^*]$, such that $(\underline{x}^*, \overline{x}^*)$ is an equilibrium for the embedding system (22), is a forward invariant set for (2). Additionally, Theorem 4 and Proposition 1 describe a computationally tractable algorithm for calculating $[\underline{x}^*, \overline{x}^*]$, which we demonstrate in the Case Studies below. Thus, Theorem 3, Theorem 4, Proposition 1, and Corollary 1 solve Problem 2.

### C. Refinement of Reachable and Forward Invariant Sets

Since we consider systems in continuous time, we can further refine the estimation of our reachable and forward invariant sets by noting that, when calculating these sets, we are only interested in the potential behavior of the disturbance at the boundary of the set. Thus, we can modify the range over which we calculate the bounds $\underline{\gamma}, \overline{\gamma}$ in our formulations to reflect this, at the cost of increased computational complexity. This idea is formalized in our next set of theoretical results which refine Theorems 1 and 3. We use the notation $a_{[i:b]}$ to denote a vector $a$ whose $i$th term has been replaced by the $i$th term of vector $b$.

**Theorem 5.** *Consider the hypotheses of Theorem 1 and construct the system*

$$\begin{bmatrix} \dot{x}' \\ \dot{\widehat{x}}' \end{bmatrix} = e'(x, \widehat{x}) := \begin{bmatrix} \delta_1(x, \underline{\gamma}(x, \widehat{x}_{[1:x]}), \widehat{x}, \overline{\gamma}(x, \widehat{x}_{[1:x]})) \\ \vdots \\ \delta_n(x, \underline{\gamma}(x, \widehat{x}_{[n:x]}), \widehat{x}, \overline{\gamma}(x, \widehat{x}_{[n:x]})) \\ \delta_1(\widehat{x}, \overline{\gamma}(x_{[1:\widehat{x}]}, \widehat{x}), x, \underline{\gamma}(x_{[1:\widehat{x}]}, \widehat{x})) \\ \vdots \\ \delta_n(\widehat{x}, \overline{\gamma}(x_{[n:\widehat{x}]}, \widehat{x}), x, \underline{\gamma}(x_{[n:\widehat{x}]}, \widehat{x})) \end{bmatrix} \tag{38}$$

*as an alternative to the embedding system* (7)*. Then,* (38) *is such that the conclusion of Theorem 1 remains valid using instead trajectories of* (38)*, that is,* $R(T, X_0) \subseteq [x'(T), \widehat{x}'(T)]$*, where* $(x'(t), \widehat{x}'(t))$ *is the solution to the* $2n$*-dimensional system* (38) *with initial condition* $(x'(0), \widehat{x}'(0)) = (\underline{x}_0, \overline{x}_0)$*.*

In Theorem 5, when determining a specific state's update behavior, we limit the calculation of the disturbance bounds to the relevant face of the state hyperrectangle, as disturbances outside of this face do not affect the state.

*Proof of Theorem 5:*
Per the problem setup and Assumption 2, the system (2) is mixed monotone with respect to a function $\delta(x, w, \widehat{x}, \widehat{w})$, where the true behavior of $w$ is dictated by an unknown function $g(x)$. Define

$$\underline{G}(\underline{x}, \overline{x}) = \max z \text{ s.t. } z \leq g(x) \forall x \in [\underline{x}, \overline{x}] \tag{39}$$

$$\overline{G}(\underline{x}, \overline{x}) = \min z \text{ s.t. } z \geq g(x) \forall x \in [\underline{x}, \overline{x}] \tag{40}$$

for $\underline{x} \preceq \overline{x}$, which results in the tightest possible inclusion function, i.e. $g(x) \in [\underline{G}(\underline{x}, \overline{x}), \overline{G}(\underline{x}, \overline{x})]$ for all $x \in [\underline{x}, \overline{x}]$.

Define a new system $\dot{x} = f'(x, W) := f(x, g(x) + W)$. Each component $f'_j(x, W)$ is mixed-monotone with respect to the decomposition function

$$\delta'_j(x, W, \widehat{x}, \widehat{W}) :=$$
$$\begin{cases} \delta_j(x, \underline{G}(x, \widehat{x}_{[j:x]}) + w^j, \widehat{x}, \overline{G}(x, \widehat{x}_{[j:x]}) + \widehat{w}^j), & x \preceq \widehat{x} \\ \delta_j(\widehat{x}, \overline{G}(x_{[j:\widehat{x}]}, \widehat{x}) + \widehat{w}^j, x, \underline{G}(x_{[j:\widehat{x}]}, \widehat{x}) + w^j), & x \succeq \widehat{x} \end{cases}$$

where $W = [w^1, w^2, ..., w^n]^T \in \mathbb{R}^{pn}$ and $\widehat{W} = [\widehat{w}^1, \widehat{w}^2, ..., \widehat{w}^n]^T \in \mathbb{R}^{pn}$. One can verify the derivative conditions for this by utilizing the chain rule and noting the following properties of $\underline{G}$ and $\overline{G}$:

$$\frac{\partial \underline{G}(x, \widehat{x}_{[i:x]})}{\partial x_j} \geq 0, \frac{\partial \overline{G}(x, \widehat{x}_{[i:x]})}{\partial x_j} \leq 0 \ \forall_{i,j \in n, i \neq j}$$

$$\frac{\partial \underline{G}(x, \widehat{x}_{[i:x]})}{\partial \widehat{x}_j} \leq 0, \frac{\partial \overline{G}(x, \widehat{x}_{[i:x]})}{\partial \widehat{x}_j} \geq 0 \ \forall_{i,j \in n}$$

$$\frac{\partial \underline{G}(x_{[i:\widehat{x}]}, \widehat{x})}{\partial x_j} \geq 0, \frac{\partial \overline{G}(x_{[i:\widehat{x}]}, \widehat{x})}{\partial x_j} \leq 0 \ \forall_{i,j \in n}$$

$$\frac{\partial \underline{G}(x_{[i:\widehat{x}]}, \widehat{x})}{\partial \widehat{x}_j} \leq 0, \frac{\partial \overline{G}(x_{[i:\widehat{x}]}, \widehat{x})}{\partial \widehat{x}_j} \geq 0 \ \forall_{i,j \in n, i \neq j}.$$

Then, construct the embedding system

$$\begin{bmatrix}\dot{x}\\ \dot{\widehat{x}}\end{bmatrix} = e(x,\widehat{x},W_1,\widehat{W}_1,W_2,\widehat{W}_2) = \begin{bmatrix}\delta'(x,W_1,\widehat{x},\widehat{W}_1)\\ \delta'(\widehat{x},\widehat{W}_2,x,W_2)\end{bmatrix} \tag{41}$$

which has inputs $(W_1,\widehat{W}_1,\widehat{W}_2,W_2) \in \mathbb{R}^{4pn}$. This is a modified embedding system without symmetric disturbances that is a monotone control system with respect to the southeast order on both state and all inputs $(w_1^1,\widehat{w}_1^1,\widehat{w}_2^1,w_2^1),...,(w_1^n,\widehat{w}_1^n,\widehat{w}_2^n,w_2^n)$.

Observe that $\underline{G}(x,x) = \overline{G}(x,x) = g(x)$. Then, with $W = \widehat{W} = 0$ and initial state $x(0) = x_0$, solutions to $\dot{x} = f'(x,0)$ are also solutions to $\dot{x} = f(x,g(x))$. Additionally, (41) consists of two copies of this true solution when $x(0) = \widehat{x}(0) = x_0$. Thus, given some $W_1,W_2 \preceq 0$ and $\widehat{W}_1,\widehat{W}_2 \succeq 0$, we can apply the fundamental result of mixed monotone systems theory (6) for solutions of (41). Namely, that the hyperrectangle defined by the trajectory $\Phi^e(t;\underline{x},\overline{x},W_1,\widehat{W}_1,W_2,\widehat{W}_2)$ of (41) overapproximates the true reachable set of the system, i.e.

$$\phi(t,x_0) \in [\![\Phi^e(t;\underline{x},\overline{x},W_1,\widehat{W}_1,W_2,\widehat{W}_2)]\!] \tag{42}$$

for $x_0 \in [\underline{x},\overline{x}]$.

Now, consider $(x'(t),\widehat{x}'(t))$ as the solution to (38) for any $(x'(0),\widehat{x}'(0)) = (\underline{x}_0,\overline{x}_0)$. By the definition of $\underline{G}$ and $\overline{G}$, there exist some $W_1,W_2 \preceq 0$ and $\widehat{W}_1,\widehat{W}_2 \succeq 0$ such that

$$\begin{aligned}\underline{\gamma}(x,\widehat{x}_{[i:x]}) &= \underline{G}(x,\widehat{x}_{[i:x]}) + w_1^i(t)\\ \overline{\gamma}(x,\widehat{x}_{[i:x]}) &= \overline{G}(x,\widehat{x}_{[i:x]}) + \widehat{w}_1^i(t)\\ \underline{\gamma}(x_{[i:\widehat{x}]},\widehat{x}) &= \underline{G}(x_{[i:\widehat{x}]},\widehat{x}) + w_2^i(t)\\ \overline{\gamma}(x_{[i:\widehat{x}]},\widehat{x}) &= \overline{G}(x_{[i:\widehat{x}]},\widehat{x}) + \widehat{w}_2^i(t)\end{aligned}$$

for all $t$. Thus, solutions to (38) are equivalent to solutions of (41) with the associated $W_1,\widehat{W}_1,W_2,\widehat{W}_2$ and therefore overapproximate the true reachable set of the system, i.e.,

$$R(T,X_0) \subseteq [\![\Phi^{e'}(T;(\underline{x}_0,\overline{x}_0))]\!] = [x'(T),\widehat{x}'(T)]. \tag{43}$$

■

Per the natural inclusion requirement in Assumption 1, modifying the disturbance bound calculation in this way produces tighter bounds, which in turn results in tighter reachable set estimates via Theorem 2. We can similarly apply these modified disturbance bounds toward the verification of forward invariant sets as outlined in Theorem 6.

**Theorem 6.** *Consider the hypotheses of Theorem 3, and suppose* $\underline{x}^*,\overline{x}^* \in \mathbb{R}^n, \underline{w}_j^{i*},\overline{w}_j^{i*} \in \mathbb{R}^p, \underline{x}^* \preceq \overline{x}^*$ *and* $\underline{w}_j^{i*} \preceq \overline{w}_j^{i*}$ *satisfy, for all* $i \in \{1,...,n\}$*,* $j \in \{1,2\}$*,*

$$\underline{w}_1^{i*} \preceq \underline{\gamma}(\underline{x}^*,\overline{x}^*_{[i:\underline{x}^*]}) \text{ and } \overline{\gamma}(\underline{x}^*,\overline{x}^*_{[i:\underline{x}^*]}) \preceq \overline{w}_1^{i*} \tag{44}$$

$$\underline{w}_2^{i*} \preceq \underline{\gamma}(\underline{x}^*_{[i:\overline{x}^*]},\overline{x}^*) \text{ and } \overline{\gamma}(\underline{x}^*_{[i:\overline{x}^*]},\overline{x}^*) \preceq \overline{w}_2^{i*} \tag{45}$$

*and*

$$\delta_i(\underline{x}^*,\underline{w}_1^{i*},\overline{x}^*,\overline{w}_1^{i*}) \succeq 0,\ \delta_i(\overline{x}^*,\overline{w}_2^{i*},\underline{x}^*,\underline{w}_2^{i*}) \preceq 0 \tag{46}$$

*as alternatives to* (20) *and* (21)*. Then the conclusion of Theorem 3 remains valid, i.e.,* $[\underline{x}^*,\overline{x}^*]$ *is forward invariant for* (2)*.*

The proof is similar to that of Theorem 3 and is omitted.

Thus far, we have considered cases where tightening the confidence bounds on the unknown component of the system results in tighter overapproximations of reachable and forward invariant sets. However, we can approach this problem from a different perspective. Regarding the challenge of safety, we can instead consider our forward-invariant sets to be *safety sets*, and attempt to expand these sets as more observations about the unknown components of the system are made. Traditionally, when approaching the problem of safety set expansion, the entire safety set is considered for observation, with expansion theoretically guaranteed by increasing confidence of the unknown behavior over the entire range [18], [25]. However, Theorem 6 allows us to guide our observation selection; to expand the safety set, we only need to update our knowledge of the unknown component of the dynamics near the edge of the current safety set. We showcase this behavior in Section VI.

In the next section, we derive an approach for calculating the bounds $\underline{\gamma},\overline{\gamma}$ such that they satisfy Assumption 1 with high probability (specifically, with probability $1-\eta$ for some small $\eta \in (0,1)$). The results presented in this section directly inherit this high-probability property, i.e., all calculated reachable sets (Theorems 1 and 2), refined reachable sets (Theorem 5), forward invariant sets (Theorems 3 and 4), and refined forward invariant sets (Theorem 6) are correct with probability at least $1-\eta$.

## V. Gaussian Processes for High Probability Uncertainty Bounds

We now propose an approach to construct functions $\underline{\gamma}$ and $\overline{\gamma}$ satisfying Assumption 1 with probability $1-\eta$. The crux of the approach is to posit the existence of a specific probability distribution over the function space for each of the unknown functions $\{g_i\}_{i=1}^p$ constituting the unknown part of the dynamics.

**Assumption 3.** *The unknown functions* $\{g_i\}_{i=1}^p$*, are independent realizations of a Gaussian Process* $\mathcal{GP}(0,k(x,x'))$ *with zero mean and kernel* $k(\cdot,\cdot)$*. In addition, observations of the GP are perturbed by additive i.i.d. Gaussian noise* $\mathcal{N}(0,\sigma^2)$ *with zero mean and variance* $\sigma^2$*.*

GPs are a form of non-parametric estimators [9] that are extremely powerful and have gained popularity in a broad range of applications including optimization and control [10]–[14], [26]. The kernel $k(\cdot,\cdot)$ is a hyperparameter of the model that controls the correlation of the GP over its domain, which can be viewed as an assumption regarding the smoothness of the unknown function $g_i$.

Assumption 3 essentially states that the unknown function is drawn from a GP, which allows the approximation of the true unknown functions $\{g_i\}_{i=1}^p$ using surrogate functions that 1) provide lower or upper bounds for the true functions with high probability; and 2) can be refined by acquiring additional observations of the GP. Specifically, given noisy observations $\{y_j\}_{j=1}^t$ of the GPs at corresponding points $\{x_j\}_{j=1}^t$ the surrogate functions of interest to approximate $g_i$ are

$$\forall i \in \{1, \cdots, p\} \quad \begin{cases} \overline{g}_i^{(t)}(x) := \mu_t(x) + \sqrt{\beta_t}\sigma_t(x) \\ \underline{g}_i^{(t)}(x) := \mu_t(x) - \sqrt{\beta_t}\sigma_t(x) \end{cases} \tag{47}$$

where $\beta_t$ is to be specified later (see Theorem 7), $\mu_t(\cdot)$ is the posterior mean, and $\sigma_t(\cdot)$ is the posterior variance, computed according to the standard GP updates [9]:

$$\mu_t(x) := k_t(x)^T (K_t + \sigma^2 I)^{-1} y \tag{48}$$

$$k_t(x, x') := k(x, x') - k_t(x)^T (K_t + \sigma^2 I)^{-1} k_t(x') \tag{49}$$

$$\sigma_t^2(x) := k_t(x, x) \tag{50}$$

where $k_t(x) := (k(x_1, x), \cdots, k(x_t, x))$ and $K_t = [k_t(x_i, x_j)]$. Intuitively, the bounds in (47) hold with high probability and can be used to create functions $\underline{\gamma}$ and $\overline{\gamma}$ as follows. For all $i \in \{1, \cdots, p\}$ and all $t \geq 1$

$$\forall \underline{x} \preceq \overline{x} \quad \begin{cases} \underline{\gamma}_i^{(t)}(\underline{x}, \overline{x}) := \min_{x \in [\underline{x}, \overline{x}]} \underline{g}_i^{(t)}(x), \\ \overline{\gamma}_i^{(t)}(\underline{x}, \overline{x}) := \max_{x \in [\underline{x}, \overline{x}]} \overline{g}_i^{(t)}(x). \end{cases} \tag{51}$$

By definition, the functions in (51) satisfy the natural inclusion property in Assumption 1. However, establishing that (3) holds requires a bit more care because the statement has to hold for all states in a subset of $\mathbb{R}^n$ and for all times $t$ at which updates are made to the GP. We introduce the following mild technical assumptions [26].

**Assumption 4.** *The states $x$ are confined to a compact subset $\mathcal{D} \subset \mathbb{R}^n$ included in a hypercube of edge size $r$. In addition, there exist constants $a, b > 0$ such that*

$$\forall i \in \{1, \cdots, n\} \forall j \in \{1, \cdots, p\} \quad \Pr\left(\sup_{x \in \mathcal{D}} \left|\frac{\partial g_j}{\partial x_i}\right| > L\right) \leq a e^{-L^2/b^2}. \tag{52}$$

The first part of Assumption 4 is relatively mild since we can assume that the trajectories are confined into a (possibly) large hypercube when computing safe reachable sets as in Section VI. The second part of Assumption 4 is also mild and is satisfied by many kernels of interest [26]. By adapting the proof of [26, Theorem 2], we have the following.

**Theorem 7.** *Assume that the unknown functions $\{g_i\}_{i=1}^p$ satisfy Assumptions 3 and 4. Pick $\eta \in (0, 1)$ and set*

$$\beta_t := 2 \log\left(\frac{pt^2\pi^2}{3\eta}\right) + 2n \log\left(t^2 nbr \sqrt{\log\left(\frac{2pna}{\eta}\right)}\right). \tag{53}$$

*At every step $t$ of the GP update, define a uniform discretization $\mathcal{D}_t$ of the hypercube containing $\mathcal{D}$ with size $\tau_t^n$ where $\tau_t := nt^2 br \sqrt{\log\left(\frac{2npa}{\eta}\right)}$. For every $x \in \mathcal{D}$, define*

$$x^{(t,-)} := \sup\{y \in \mathcal{D}_t \mid y \preceq x\}, \tag{54}$$

$$x^{(t,+)} := \inf\{y \in \mathcal{D}_t \mid x \preceq y\}. \tag{55}$$

*For all $i \in \{1, \cdots, p\}$ and all $t \geq 1$ define $\forall \underline{x} \preceq \overline{x}$*

$$\underline{\gamma}_i^{(t)}(\underline{x}, \overline{x}) := \min_{x \in [\underline{x}^{(t,-)}, \overline{x}^{(t,+)}] \cap \mathcal{D}_t} \underline{g}_i^{(t)}(x) - \frac{1}{t^2}, \tag{56}$$

$$\overline{\gamma}_i^{(t)}(\underline{x}, \overline{x}) := \max_{x \in [\underline{x}^{(t,-)}, \overline{x}^{(t,+)}] \cap \mathcal{D}_t} \overline{g}_i^{(t)}(x) + \frac{1}{t^2}. \tag{57}$$

*Then, with probability at least $1 - \eta$,*

$$\forall \underline{x} \preceq \overline{x} \quad \forall x \in [\underline{x}, \overline{x}] \quad \forall t \geq 1 \quad \forall j \in \{1, \cdots, p\} \quad \underline{\gamma}_i^{(t)}(\underline{x}, \overline{x}) \leq g_i(x) \leq \overline{\gamma}_i^{(t)}(\underline{x}, \overline{x}). \tag{58}$$

*Sketch of proof:*

Our proof closely follows the analysis of [27, Appendix I, Section 2)], which details the result of [26, Theorem 2]. The idea is to combine the finite discretization $\mathcal{D}_t$ of $\mathcal{D}$ at each step $t$, for which the GP approximations hold, with Assumption 4 to bound the maximum deviation for points outside the discretization.

With the choice of $\tau_t$ and $|\mathcal{D}_t| = \tau_t^n$, the choice of $\beta_t$ ensures [27, Lemma 5.6] that with probability at least $1-\eta/2$

$$\forall x \in \mathcal{D}_t \quad \forall i \in \{1, \ldots, p\} \quad \forall t \geq 1 \quad \underline{g}_i^{(t)}(x) \leq g_i(x) \leq \overline{g}_i^{(t)}(x). \tag{59}$$

Set $\tau := b\sqrt{\log \frac{2npa}{\eta}}$ and note that $\frac{\eta}{2} = npa \exp\left(-\frac{\tau^2}{b^2}\right)$. From Assumption 4, with probability at least $1 - \eta/2$

$$\forall i \in \{1, \ldots, n\} \quad \forall j \in \{1, \ldots, p\} \quad \sup_{x \in \mathcal{D}} \left|\frac{\partial g_i}{\partial x_j}\right| \leq \tau. \tag{60}$$

From (60), we have with probability at least $1 - \eta/2$

$$\forall x, x' \in \mathcal{D} \quad \forall i \in \{1, \ldots, n\} \quad \forall j \in \{1, \ldots, p\} \quad |g_i(x) - g_i(x')| \leq \tau \|x - x'\|_1. \tag{61}$$

When obtaining $\mathcal{D}_t$ as a uniform discretization of size $\tau_t^n$ of the hypercube containing $\mathcal{D}$, we then have

$$\forall x \in \mathcal{D} \quad \|x - x^{(t,-)}\|_1 \leq \frac{nr}{\tau_t} \text{ and } \|x - x^{(t,+)}\|_1 \leq \frac{nr}{\tau_t}.$$

Consequently, with probability at least $1 - \eta/2$ we have

$$\forall x \in \mathcal{D} \quad \forall i \in \{1, \ldots, n\} \quad \forall j \in \{1, \ldots, p\} \quad \begin{cases} \left|g_i(x) - g_i(x^{(t,+)})\right| \leq \tau \frac{nr}{\tau_t} = \frac{1}{t^2}, \\ \left|g_i(x) - g_i(x^{(t,-)})\right| \leq \tau \frac{nr}{\tau_t} = \frac{1}{t^2}. \end{cases} \tag{62}$$

Combining (59) and (62), with probability at least $1 - \eta$,

$$\forall x \in \mathcal{D} \quad \forall i \in \{1, \ldots, n\} \quad \forall j \in \{1, \ldots, p\} \quad \begin{cases} g_i(x) \leq \overline{g}_i^{(t)}(x^{(t,+)}) + \frac{1}{t^2}, \\ g_i(x) \geq \underline{g}_i^{(t)}(x^{(t,-)}) - \frac{1}{t^2}, \end{cases} \tag{63}$$

Consequently, with probability at least $1 - \eta$, $\forall \underline{x} \preceq \overline{x}$, $\forall x \in [\underline{x}; \overline{x}]$, $\forall j \in \{1, \ldots, p\}$ and $\forall t \geq 1$

$$\underline{\gamma}_i^{(t)}(\underline{x}, \overline{x}) := \min_{x' \in [\underline{x}^{(t,-)}, \overline{x}^{(t,+)}] \cap \mathcal{D}_t} \underline{g}_i^{(t)}(x') - \frac{1}{t^2} \tag{64}$$

$$\leq \min_{x' \in [\underline{x}^{(t,-)}, \overline{x}^{(t,+)}]} g_i(x') \tag{65}$$

$$\leq g_i(x) \tag{66}$$

$$\leq \max_{x' \in [\underline{x}^{(t,-)}, \overline{x}^{(t,+)}]} g_i(x') \tag{67}$$

$$\leq \max_{x' \in [\underline{x}^{(t,-)}, \overline{x}^{(t,+)}] \cap \mathcal{D}_t} \overline{g}_i^{(t)}(x') + \frac{1}{t^2} \tag{68}$$

$$:= \overline{\gamma}_i^{(t)}(\underline{x}, \overline{x}), \tag{69}$$

which is the desired result. ■

The bounds (56) and (57) of Theorem 7 can be directly inserted into the formulations outlined in Section IV to calculate reachable and forward invariant sets that hold with probability at least $1 - \eta$.

## VI. Case Studies

We first present a case study of a vehicle encountering hazardous road conditions, illustrating the learning capabilities outlined by Theorems 1 and 2. We then provide an academic example of the algorithm outlined by Theorem 4 and Proposition 1. Finally, we showcase a case study of a multirotor expanding its known safety set in an unknown wind field by taking advantage of Theorem 6. Another example of a multirotor iteratively exploring a wind field by computing safe reachable sets via Theorem 1 is available in the conference paper [21]. Additionally, a code repository is available at `https://github.com/gtfactslab/Cao_OJCSYS2022`.

### A. Case Study: Vehicle on an Icy Road

Even though the results above focus on the time-invariant system (2) for notational simplicity, mixed monotonicity, and by extension our results, can naturally accommodate systems subject to known, time-varying, exogenously defined inputs for the reachable set overapproximation, as we demonstrate in this case study.

We consider the kinematic planar bicycle model [28] for abstracting the dynamics of a vehicle. Under nominal conditions, the model relates the positional coordinates $X$ and $Y$, center-of-mass velocity $v$, heading angle $\psi$, side-slip angle $\beta(u_2)$, and front and rear distances from center of mass $l_f$ and $l_r$ as

$$\dot{X} = v\cos(\psi + \beta(u_2)), \;\; \dot{Y} = v\sin(\psi + \beta(u_2)),$$
$$\dot{\psi} = \frac{v}{l_r}\sin(\beta(u_2)), \;\; \dot{v} = u_1, \tag{70}$$

where

$$\beta(u_2) = \arctan\left(\frac{l_r}{l_f + l_r}\tan(u_2)\right), \tag{71}$$

with inputs to the system being the desired acceleration $u_1$ and steering angle $u_2$. This model is visualized in Figure 1.

We consider a scenario where this system is suddenly subject to road conditions that affect the friction between the tires and the road surface, resulting in a disturbance in the velocity dynamics in (70) so that the resulting velocity update dynamics are given by

$$\dot{v}_{\text{actual}} = u_1 + g(u_1), \tag{72}$$

where $g(u_1)$ constitutes the unknown effect of the changed friction coefficient between the tires and road.

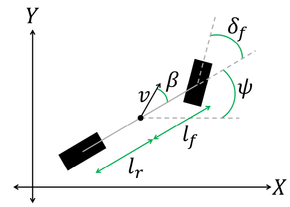


**FIGURE 1.** The planar bicycle model has positions $X$ and $Y$, center-of-mass velocity $v$, heading angle $\psi$, side-slip angle $\beta(\delta_f)$, front and rear distances from center of mass $l_f$ and $l_r$. The inputs to the system are the desired acceleration $u_1$ and front steering angle $u_2 = \delta_f$.

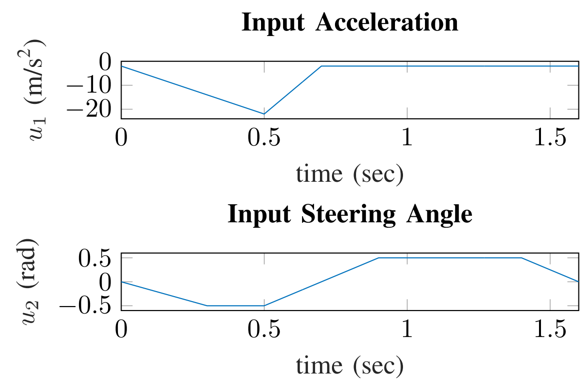


**FIGURE 2.** Known acceleration and steering angle inputs to the vehicle.

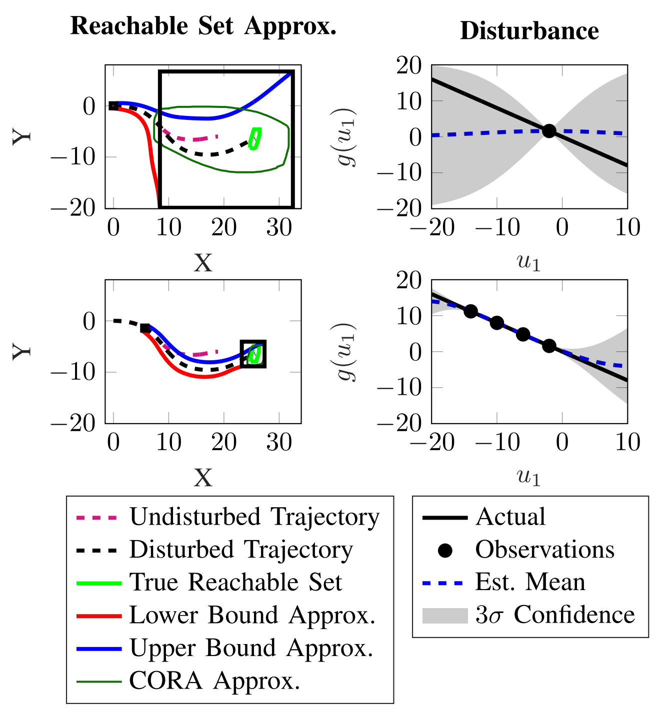


**FIGURE 3.** Visualization of the hyperrectangular overapproximation of forward reachable set using mixed monotonicity in the autonomous vehicle case study. The disturbance affecting the system is such that the true trajectory of the system notably differs from the undisturbed trajectory (left column, dotted lines). Initially, the system has low knowledge of the disturbance behavior, resulting in a conservative overapproximation of the true reachable set of the vehicle (top row). After obtaining several observations, the reachable set overapproximation is able to tightly approximate the true reachable set (bottom row).

We assume the inputs $u_1(t)$ and $u_2(t)$ follow the fixed braking and turning maneuver shown in Figure 2. We also assume uncertainty in the initial state of the system such that the initial $X$ and $Y$ positions are accurate within $0.5$m, and the heading angle is accurate within $0.05$rad. The change in the system dynamics (i.e. the road becoming slippery) occurs at time $t = 0$. As shown by the dashed lines in the left plots of Figure 3, this disturbance behavior is enough to cause the actual position of the vehicle to be notably different by the end of the $1.6$s maneuver.

A decomposition function for the planar bicycle model with added unknown disturbance (70)-(72) is given as follows. To accommodate the inputs, the associated terms in the decomposition function must simply adhere to requirement 4 in Assumption 2, i.e. $\frac{\partial \delta_i}{\partial u_k}(x, u, w, \widehat{x}, \widehat{u}, \widehat{w}) \geq 0$ and $\frac{\partial \delta_i}{\partial \widehat{u}_k}(x, u, w, \widehat{x}, \widehat{u}, \widehat{w}) \leq 0$. We take

$$\delta(x, u, w, \widehat{x}, \widehat{u}, \widehat{w}) = \begin{bmatrix} d^X \\ d^Y \\ d^\psi \\ d^v \end{bmatrix} \tag{73}$$

where

$$\begin{aligned} d^X &= d^{b_1 b_2}\Bigg( \begin{bmatrix} v \\ d^{\cos}(\psi + \beta(\delta_f), \widehat{\psi} + \beta(\widehat{\delta}_f)) \end{bmatrix}, \\ &\qquad \begin{bmatrix} \widehat{v} \\ d^{\cos}(\widehat{\psi} + \beta(\widehat{\delta}_f), \psi + \beta(\delta_f)) \end{bmatrix} \Bigg) \\ d^Y &= d^{b_1 b_2}\Bigg( \begin{bmatrix} v \\ d^{\sin}(\psi + \beta(\delta_f), \widehat{\psi} + \beta(\widehat{\delta}_f)) \end{bmatrix}, \\ &\qquad \begin{bmatrix} \widehat{v} \\ d^{\sin}(\widehat{\psi} + \beta(\widehat{\delta}_f), \psi + \beta(\delta_f)) \end{bmatrix} \Bigg) \\ d^\psi &= d^{b_1 b_2}\Bigg( \begin{bmatrix} v \\ d^{\sin}(\beta(\delta_f), \beta(\widehat{\delta}_f)) \end{bmatrix}, \begin{bmatrix} \widehat{v} \\ d^{\sin}(\beta(\widehat{\delta}_f), \beta(\delta_f)) \end{bmatrix} \Bigg) \\ d^v &= u_1 + w, \end{aligned}$$

where, for vectors $b, \widehat{b} \in \mathbb{R}^2$,

$$d^{b_1 b_2}(b, \widehat{b}) = \begin{cases} \min\{b_1 b_2, \widehat{b}_1 b_2, b_1 \widehat{b}_2, \widehat{b}_1 \widehat{b}_2\}, & \text{if } b \preceq \widehat{b} \\ \max\{b_1 b_2, \widehat{b}_1 b_2, b_1 \widehat{b}_2, \widehat{b}_1 \widehat{b}_2\}, & \text{if } \widehat{b} \preceq b, \end{cases}$$

and $d^{\sin}$, $d^{\cos}$ are tight decomposition functions for the respective trigonometric functions, which take the forms of equations (74) and (75).

By modeling $g(u)$ as a GP, we attain high-confidence bounds $\underline{\gamma}(u, \widehat{u}), \overline{\gamma}(u, \widehat{u})$ on the behavior of $g(u)$ which can be inserted into (73). Furthermore, since the input $u$ is known (i.e. $u = \widehat{u}$), the computations of $\underline{\gamma}, \overline{\gamma}$ along the system trajectory are simplified. For this example, the $\pm 3\sigma$ bounds of the GP are used. This decomposition function is then used to create the embedding system which provides hyperrectangular overapproximations of the reachable set of the system at the end of the maneuver using Theorem 1. Every $0.1$ seconds throughout the maneuver, an observation is collected about the disturbance's behavior, allowing for improved reachable set overapproximation from the updated embedding system.

$$d^{\sin}(x, \widehat{x}) := \tag{74}$$
$$\begin{cases} \sin(x), & \text{if } (\cos(x), \cos(\widehat{x})) \succeq 0 \\ & \text{ and } |x - \widehat{x}| \leq \pi \\ \sin(\widehat{x}), & \text{if } (\cos(x), \cos(\widehat{x})) \preceq 0 \\ & \text{ and } |x - \widehat{x}| \leq \pi \\ \text{sign}(x - \widehat{x}), & \text{if } |x - \widehat{x}| \geq 2\pi \\ \text{sign}(x - \widehat{x}), & \text{if } \cos(x) \leq 0 \leq \cos(\widehat{x}) \\ & \text{ and } |x - \widehat{x}| \leq 2\pi \\ \text{sign}(x - \widehat{x}), & \text{if } \cos(x)\cos(\widehat{x}) \geq 0 \\ & \text{ and } \pi \leq |x - \widehat{x}| \leq 2\pi \\ \min\{\sin(x), \sin(\widehat{x})\}, & \text{if } x \leq \widehat{x} \\ & \text{ and } \cos(x) \geq 0 \geq \cos(\widehat{x}) \\ & \text{ and } |x - \widehat{x}| \leq 2\pi \\ \max\{\sin(x), \sin(\widehat{x})\}, & \text{if } x \geq \widehat{x} \\ & \text{ and } \cos(x) \geq 0 \geq \cos(\widehat{x}) \\ & \text{ and } |x - \widehat{x}| \leq 2\pi, \end{cases}$$

$$d^{\cos}(x, \widehat{x}) := d^{\sin}\left( x + \frac{\pi}{2}, \widehat{x} + \frac{\pi}{2} \right) \tag{75}$$

To serve as a baseline, we leverage the Level Set Toolbox [29] and CORA [30] to provide approximations of the reachable set. The state-dependent disturbance behavior can be added in the Level Set Toolbox by including the calculation of the worst-case disturbance bound in the Hamiltonian and partial functions. However, due to the large state space as well as the change in scale from the initial set to the final set, the Level Set Toolbox takes several hours to return a result, and moreover this result was inaccurate, most likely due to numerical issues. Out of the box, CORA does not allow disturbance bounds to evolve in the state-dependent way as described by the problem, and we are unaware of any other toolboxes that can do so. In CORA, the state-dependent disturbance can be approximated by iteratively solving for the reachable set over a small timestep and inserting the worst-case bounds that will be found during that timestep.

As shown in Figure 3, the initial overapproximation of the reachable set is conservative due to having little knowledge about the disturbance behavior. However, after $0.3$ seconds, the embedding system is able to provide a much tighter approximation of the reachable set. Additionally, the update of the reachable set overapproximation takes approximately 0.02 seconds on a personal computer with prototype code written in MATLAB, demonstrating the potential for real-time deployment. By contrast, the set approximations given by CORA are less conservative, but take around 600 seconds (10 minutes) to compute on the same machine.

### *B. Example: Equilibrium Sequence Convergence*

We illustrate Theorem 4 and Proposition 1 via an academic example. Consider the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = F(x, w) = \begin{bmatrix} -x_1 - x_1^3 - x_2 - w \\ -x_2 - x_2^3 + x_1 + w^3 \end{bmatrix} \tag{76}$$

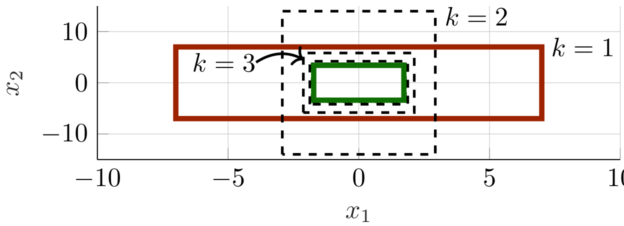


**FIGURE 4.** **Demonstration of Theorem 4 using embedding system** (80) **and disturbance bounds** (77) **and** (78), $m = 2$. $[\underline{x}^1, \overline{x}^1]$ **is initialized to** $[(-7, -7), (7, 7)]$**, shown in red (outermost solid), and the subsequent steps in the sequence are plotted in black (dashed), with the final converged hyperrectangle shown in green (innermost solid). At** $k = 2$**, the sequence fulfills** (26)**, from which we conclude that the remaining subsequence is ordered and converges from the second part of Theorem 4. The resulting hyperrectangle,** $[(-1.72, -3.40), (1.72, 3.40)]$ **is an equilibrium of** (80) **and is thus forward invariant for** (76) **per Theorem 4.**

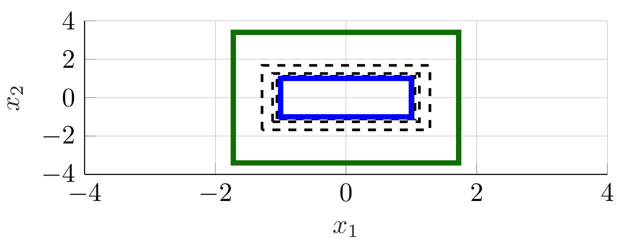


**FIGURE 5.** **Demonstration of Proposition 1 using embedding system** (80) **and disturbance bounds** (77) **and** (78), $m = 1$. **Note that** $(\underline{\gamma}, \overline{\gamma}) = (\underline{\gamma}(\underline{x}, \overline{x}; 1), \overline{\gamma}(\underline{x}, \overline{x}; 1))$ **and** $(\underline{\gamma}', \overline{\gamma}') = (\underline{\gamma}(\underline{x}, \overline{x}; 2), \overline{\gamma}(\underline{x}, \overline{x}; 2))$ **fulfill** (33). $[\underline{x}^1, \overline{x}^1]$ **is initialized to the result from the first scenario,** $[(-1.72, -3.40), (1.72, 3.40)]$**, shown in green (outermost solid), and the subsequent steps in the sequence are plotted in black (dashed), with the final converged hyperrectangle shown in blue (innermost solid). The resulting hyperrectangle,** $[(-1.00, -1.00), (1.00, 1.00)]$ **is an equilibrium of** (80) **and is thus forward invariant for** (76) **per Theorem 4.**

with state $x = (x_1, x_2) \in \mathbb{R}^2$ and disturbance $w \in \mathbb{R}$. We assume that $w$ is bounded by the parameterized functions $\underline{\gamma}(\underline{x}, \overline{x}; m) \leq w \leq \overline{\gamma}(\underline{x}, \overline{x}; m)$,

$$\underline{\gamma}(\underline{x}, \overline{x}; m) = -m(\max\{|\underline{x}_1|, |\overline{x}_1|\}) \tag{77}$$

$$\overline{\gamma}(\underline{x}, \overline{x}; m) = m(\max\{|\underline{x}_1|, |\overline{x}_1|\}) \tag{78}$$

where $m$ is a slope parameter.

The system (76) is mixed monotone with respect to the decomposition function

$$\delta(x, w, \widehat{x}, \widehat{w}) = \begin{bmatrix} -x_1 - x_1^3 - \widehat{x}_2 - \widehat{w} \\ -x_2 - x_2^3 + x_1 + w^3 \end{bmatrix}. \tag{79}$$

Thus, we form the associated embedding system using (7),

$$\begin{bmatrix} \dot{x} \\ \dot{\widehat{x}} \end{bmatrix} = e(x, \widehat{x}) := \begin{bmatrix} -x_1 - x_1^3 - \widehat{x}_2 - \overline{\gamma}(x, \widehat{x}; m) \\ -x_2 - x_2^3 + x_1 + \underline{\gamma}(x, \widehat{x}; m)^3 \\ -\widehat{x}_1 - \widehat{x}_1^3 - x_2 - \underline{\gamma}(x, \widehat{x}; m) \\ -\widehat{x}_2 - \widehat{x}_2^3 + \widehat{x}_1 + \overline{\gamma}(x, \widehat{x}; m)^3 \end{bmatrix}. \tag{80}$$

We then consider two scenarios in which a forward invariant set of (76) is calculated by determining an equilibrium of the system. In the first scenario we take $m = 2$, then in the second scenario we initialize the equilibrium sequence to the results of the first scenario and take $m = 1$. These scenarios are demonstrated in Figures 4 and 5, respectively.

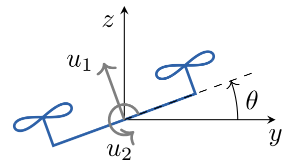


**FIGURE 6.** **The planar multirotor model has horizontal position** $y$**, vertical position** $z$**, and roll angle** $\theta$**. The inputs are thrust** $u_1$ **in the direction perpendicular to the line segment connecting the rotors and roll angle velocity** $u_2$**. The five dimensional state** $x$ **consists of** $y$, $z$, $\theta$**, and the derivatives** $v_y = \dot{y}$, $v_z = \dot{z}$**, so that** $x = [y \; v_y \; z \; v_z \; \theta]^T$**.**

### *C. Case Study: Multirotor Exploring a Wind Field*

We consider a planar multirotor model, illustrated in Figure 6, for a multirotor aerial vehicle that is constrained to move in a vertical plane. The horizontal and vertical position of the multirotor are denoted $y$ and $z$, and the roll angle is denoted $\theta$. The system has two inputs, thrust $u_1$ acting at the center of mass in the direction $\begin{bmatrix} -\sin\theta & \cos\theta \end{bmatrix}^T$, and the roll angular velocity $u_2$. We also assume that the system is subject to input saturation, gravitational acceleration $a_g$, and an unknown force due to wind. This wind force affects the acceleration in both the horizontal and vertical directions, and the force in each direction is a function of its associated state (i.e. the horizontal wind force is a function of position $y$ and the vertical wind force is a function of altitude $z$). The resulting dynamics with normalized mass and moment of inertia are given by

$$\begin{aligned} \ddot{y} &= -u_1 \sin\theta + g_1(y) \\ \ddot{z} &= u_1 \cos\theta - a_g + g_2(z) \\ \dot{\theta} &= u_2 \end{aligned} \tag{81}$$

where $g_1$ and $g_2$ constitute the unknown wind forces in the horizontal and vertical directions, respectively. We estimate $g_1$ and $g_2$ using GPs with a radial basis function kernel, and obtain high confidence bounds by considering posterior estimates up to two standard deviations from the mean.

To control the system, we employ feedback linearization to transform the system into a set of triple integrators in the $y$ and $z$ directions. We then perform an eigenvector transformation and utilize the method described in [1] to derive a decomposition function for the resulting feedback linearized system. This decomposition function allows us to determine high probability forward invariant sets that account for the potential disturbance modeled by the GPs. We then develop an operational controller and a safety controller using the feedback linearized system. The operational controller utilizes proportional-integral-derivative (PID) control to drive the system to a target point, while the safety controller uses proportional-derivative (PD) control to drive the system to the origin. During operation, an invariant safe set is computed using Theorem 5 for the closed loop system with safety controller, such that the safety controller does not produce inputs that will saturate. The overall control strategy is to execute the operational controller action to explore the safety set and collect measurements, and to switch to the safety controller if the system enters within some distance of the safety set boundary in order to maintain safety with high probability.

We assume there exist obstacles which the system must avoid as shown in Figure 7. Each GP is initialized with observations near the starting point of the multirotor vehicle. Using Theorem 5 and numerically searching for states in the embedding system that satisfy Theorem 6, we find the largest invariant set that does not intersect the obstacles and thus can be used as a safety set, as well as the smallest invariant set that intersects the obstacles and therefore cannot be used as a safety set, as shown in Figure 7. We also leverage Theorems 1 and 3 and the same numerical search method to find another invariant set, which can be seen in the figure to be smaller and thus more restrictive.

We then task the system with safely collecting observations about the wind's behavior in order to update the GPs, allowing for iterative computation of new, larger safety sets. Recalling that Theorem 6 allows us to focus observation points near the edge of the safety set, we have the system fly to observation points near the corners of the current safety set (marked by the asterisks in the figures), take measurements of the disturbance, and then update the GPs and calculate a new forward invariant set. This process continues iteratively, allowing for the multirotor to safely explore larger areas of the state space with high probability guarantees of avoiding unsafe regions, as shown in Figure 7. For comparison, we use the same procedure with Theorems 1 and 3. After the same number of updates, we can again see in the figure that this produces a more restrictive invariant set. On average, the sets found by Theorem 6 took around 1.8 seconds to produce on a personal computer with code written in MATLAB, while the sets found by Theorem 3 took around 0.8 seconds, showcasing the tradeoff between these theorems.

## VII. Conclusion

We have presented a technique for efficiently computing hyperrectangular reachable and forward invariant sets for dynamical systems with unknown components by modeling these components as GPs, which allows for the assumption of high probability bounds. For systems that are moderately high dimensional in the known component and low dimensional in the unknown component, which occur often in practice, this assumption in turn enables tractable algorithms for determining reachable and forward invariant sets with high probability by leveraging tools from mixed monotone systems theory. As shown in the examples, this is achieved by sampling the GP over the relevant region of the state space to retrieve high probability bounds, then incorporating these bounds into an embedding system, from which the reachable and forward invariant sets are calculated. In contrast to prior literature, this method does not require discrete-time system dynamics or that the system dynamics be affine in the unknown components. Future directions of research include exploring sparse Gaussian process estimation to further improve the computational performance, as well as implementation on higher-dimensional or physical systems.

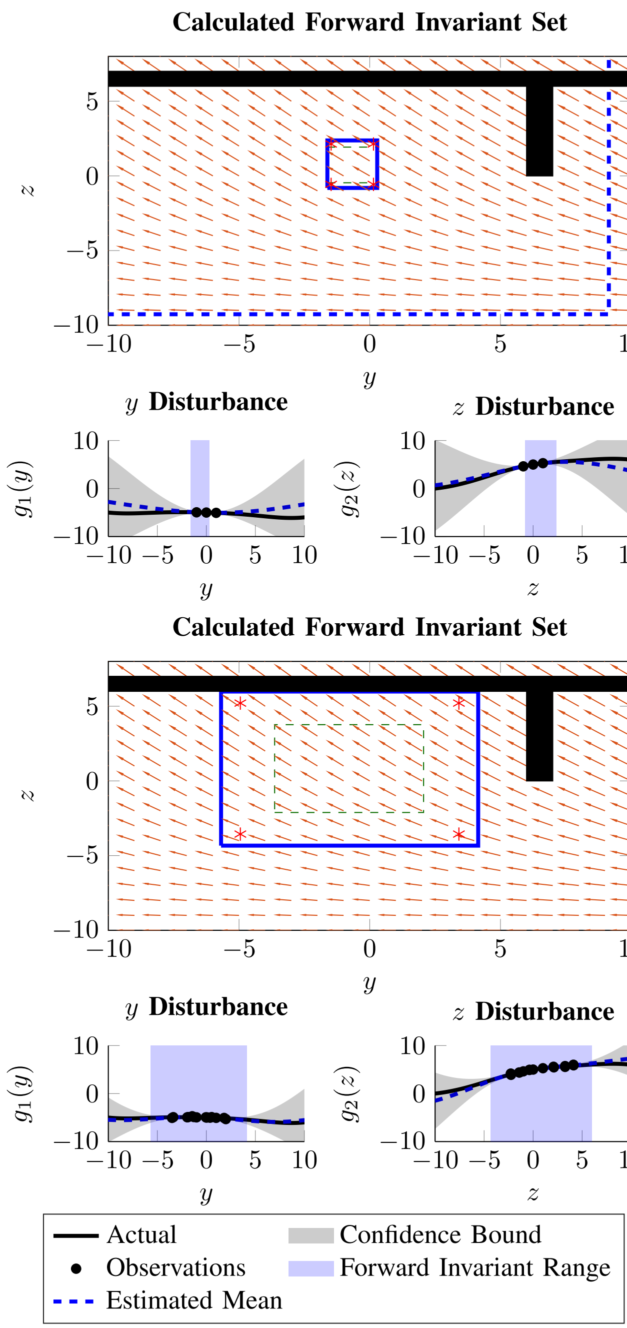


**FIGURE 7.** **The planar multirotor operates in an unknown wind field with several obstacles and is initialized with some observations (2nd row) of the wind's behavior (1st row, arrows). Using Theorem 6, we identify two invariant sets for the closed-loop safety controller (1st row, blue solid and dashed rectangles). However, only the smaller (blue solid) set can be used initially as a safety set because the larger set (blue dashed) intersects obstacles. Additionally, the largest acceptable set (1st row, green dashed) produced by Theorem 3 is more restrictive. After collecting several rounds of observations (4th row), the multirotor can now safely explore near the obstacles without collision, as denoted by the new safety set (3rd row, solid blue). Again, note that set produced by Theorem 3 (3rd row, green dashed) is more restrictive.**

## REFERENCES


[1] S. Coogan, "Mixed monotonicity for reachability and safety in dynamical systems," in *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 5074–5085, 2020.
[2] S. Coogan, E. A. Gol, M. Arcak, and C. Belta, "Traffic network control from temporal logic specifications," *IEEE Transactions on Control of Network Systems*, vol. 3, pp. 162–172, June 2016.
[3] M. Abate, M. Mote, E. Feron, and S. Coogan, "Verification and runtime assurance for dynamical systems with uncertainty," in *Hybrid Systems: Computation and Control*, 2021.
[4] M. Abate and S. Coogan, "Computing robustly forward invariant sets for mixed-monotone systems," in *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 4553–4559, 2020.
[5] G. Enciso, H. Smith, and E. Sontag, "Nonmonotone systems decomposable into monotone systems with negative feedback," *Journal of Differential Equations*, vol. 224, no. 1, pp. 205–227, 2006.
[6] D. Angeli, G. A. Enciso, and E. D. Sontag, "A small-gain result for orthant-monotone systems under mixed feedback," *Systems & Control Letters*, vol. 68, pp. 9–19, 2014.
[7] S. Coogan and M. Arcak, "Efficient finite abstraction of mixed monotone systems," in *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*, pp. 58–67, 2015.
[8] H. Smith, "Global stability for mixed monotone systems," *Journal of Difference Equations and Applications*, vol. 14, no. 10-11, pp. 1159–1164, 2008.
[9] C. E. Rasmussen and C. K. I. Williams., *Gaussian Processes for Machine Learning*. Cambridge, Massachusetts: MIT Press, 2006.
[10] F. Berkenkamp, A. P. Schoellig, and A. Krause, "Safe controller optimization for quadrotors with gaussian processes," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 491–496, 2016.
[11] M. Liu, G. Chowdhary, B. C. da Silva, S.-Y. Liu, and J. P. How, "Gaussian processes for learning and control: A tutorial with examples," *IEEE Control Systems Magazine*, vol. 38, pp. 53–86, oct 2018.
[12] W. Luo, C. Nam, G. Kantor, and K. Sycara, "Distributed environmental modeling and adaptive sampling for multi-robot sensor coverage," in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '19, (Richland, SC), pp. 1488–1496, International Foundation for Autonomous Agents and Multiagent Systems, 2019.
[13] A. Benevento, M. Santos, G. Notarstefano, K. Paynabar, M. Bloch, and M. Egerstedt, "Multi-robot coordination for estimation and coverage of unknown spatial fields," in *Proc. of IEEE International Conference on Robotics and Automation*, (Paris, France), pp. 7740–7746, May 2020.
[14] A. Lederer, J. Umlauft, and S. Hirche, "Uniform error bounds for gaussian process regression with application to safe control," in *Advances in Neural Information Processing Systems* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds.), vol. 32, (Vancouver, Canada), Curran Associates, Inc., Dec. 2019.
[15] M. J. Khojasteh, V. Dhiman, M. Franceschetti, and N. Atanasov, "Probabilistic safety constraints for learned high relative degree system dynamics," in *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, vol. 120 of *Proceedings of Machine Learning Research*, pp. 781–792, PMLR, 10–11 Jun 2020.
[16] T. Koller, F. Berkenkamp, M. Turchetta, and A. Krause, "Learning-based model predictive control for safe exploration," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 6059–6066, 2018.
[17] J. J. Choi, F. Castañeda, C. J. Tomlin, and K. Sreenath, "Reinforcement learning for safety-critical control under model uncertainty, using control lyapunov functions and control barrier functions," in *Proceedings of Robotics: Science and Systems*, (Corvalis, Oregon, USA), July 2020.
[18] F. Castañeda, J. J. Choi, B. Zhang, C. J. Tomlin, and K. Sreenath, "Gaussian process-based min-norm stabilizing controller for control-affine systems with uncertain input effects and dynamics," in *2021 American Control Conference (ACC)*, pp. 3683–3690, 2021.
[19] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, "A general safety framework for learning-based control in uncertain robotic systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 2737–2752, 2019.
[20] S. Herbert, J. J. Choi, S. Sanjeev, M. Gibson, K. Sreenath, and C. J. Tomlin, "Scalable learning of safety guarantees for autonomous systems using hamilton-jacobi reachability," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5914–5920, 2021.
[21] M. E. Cao, M. Bloch, and S. Coogan, "Estimating high probability reachable sets using gaussian processes," in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 3881–3886, 2021.
[22] M. Abate, M. Dutreix, and S. Coogan, "Tight decomposition functions for continuous-time mixed-monotone systems with disturbances," *IEEE Control Systems Letters*, vol. 5, no. 1, pp. 139–144, 2021.
[23] D. Angeli and E. D. Sontag, "Monotone control systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 10, pp. 1684–1698, 2003.
[24] H. Smith, *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*. Mathematical surveys and monographs, American Mathematical Society, 2008.
[25] F. Berkenkamp, R. Moriconi, A. P. Schoellig, and A. Krause, "Safe learning of regions of attraction for uncertain, nonlinear systems with Gaussian processes," in *Proc. of the IEEE Conference on Decision and Control (CDC)*, pp. 4661–4666, 2016.
[26] N. Srinivas, A. Krause, S. Kakade, and M. Seeger, "Gaussian process optimization in the bandit setting: No regret and experimental design," in *Proceedings of the 27th International Conference on Machine Learning*, (USA), pp. 1015–1022, 2010.
[27] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for gaussian process optimization in the bandit setting," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3250–3265, 2012.
[28] P. Polack, F. Altché, B. Novel, and A. de La Fortelle, "The kinematic bicycle model: A consistent model for planning feasible trajectories for autonomous vehicles?," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, pp. 812–818, June 2017.
[29] I. M. Mitchell and J. A. Templeton, "A toolbox of hamilton-jacobi solvers for analysis of nondeterministic continuous and hybrid systems," in *Hybrid Systems: Computation and Control* (M. Morari and L. Thiele, eds.), (Berlin, Heidelberg), pp. 480–494, Springer Berlin Heidelberg, 2005.
[30] M. Althoff, "An introduction to cora 2015," in *Proc. of the workshop on applied verification for continuous and hybrid systems*, pp. 120–151, 2015.


**MICHAEL ENQI CAO** (Student Member, IEEE) received the B.S. degree in Computer Engineering in 2018 followed by the M.S. degree in Electrical and Computer Engineering in 2020 from the Georgia Institute of Technology. He has done work for NASA's Jet Propulsion Lab and Johns Hopkins University's Applied Physics Lab, and is currently pursuing the Ph.D. degree in Robotics at the Georgia Institute of Technology.

**MATTHIEU BLOCH** (Senior Member, IEEE) is a Professor in the School of Electrical and Computer Engineering. He received the Engineering degree from Supélec, Gif-sur-Yvette, France, the M.S. degree in Electrical Engineering from the Georgia Institute of Technology, Atlanta, in 2003, the Ph.D. degree in Engineering Science from the Université de Franche-Comté, Besançon, France, in 2006, and the Ph.D. degree in Electrical Engineering from the Georgia Institute of Technology in 2008. In 2008-2009, he was a postdoctoral research associate at the University of Notre Dame, South Bend, IN. Since July 2009, Dr. Bloch has been on the faculty of the School of Electrical and Computer Engineering, and from 2009 to 2013 Dr. Bloch was based at Georgia Tech Lorraine. His research interests are in the areas of information theory, error-control coding, wireless communications, and cryptography. Dr. Bloch has served on the organizing committee of several international conferences; he was the chair of the Online Committee of the IEEE Information Theory Society from 2011 to

2014, an Associate Editor for the IEEE Transactions on Information Theory from 2016 to 2019 and again since 2021, and he has been on the Board of Governors of the IEEE Information Theory Society since 2016 and currently serves as the 1st Vice-President. He has been an Associate Editor for the IEEE Transactions on Information Forensics and Security since 2019. He is the co-recipient of the IEEE Communications Society and IEEE Information Theory Society 2011 Joint Paper Award and the co-author of the textbook Physical-Layer Security: From Information Theory to Security Engineering published by Cambridge University Press.

**SAMUEL COOGAN** (Senior Member, IEEE) received the B.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA in 2010 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, Berkeley, CA, USA in 2012 and 2015.

He is currently an associate professor and the Demetrius T. Paris Junior Professor at the Georgia Institute of Technology, Atlanta, GA, USA in the School of Electrical and Computer Engineering and the School of Civil and Environmental Engineering. Prior to joining Georgia Tech in 2017, he was an assistant professor at the University of California, Los Angeles from 2015 to 2017.

Prof. Coogan received a CAREER Award from the National Science Foundation in 2018, a Young Investigator Award from the Air Force Office of Scientific Research in 2019, and the Donald P Eckman Award from the American Automatic Control Council in 2020. He is a member of SIAM and a senior member of IEEE.