Input Design for Nonlinear Model Discrimination via Affine Abstraction ⋆

Kanishka Singh * **Yuhao Ding** ** **Necmiye Ozay** **
Sze Zheng Yong *

* *School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, 85287 (emails:* {krsingh1,szyong}@asu.edu*)*
** *Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI, 48109 (emails:* {yuhding,necmiye}@umich.edu*)*

Abstract: This paper considers the design of *separating* input signals in order to discriminate among a finite number of uncertain nonlinear models. Each nonlinear model corresponds to a system operating mode, unobserved intents of other drivers or robots, or to fault types or attack strategies, etc., and the separating inputs are designed such that the output trajectories of all the nonlinear models are guaranteed to be distinguishable from each other under any realization of uncertainties in the initial condition, model discrepancies or noise. We propose a two-step approach. First, using an optimization-based approach, we over-approximate nonlinear dynamics by uncertain affine models, as abstractions that preserve all its system behaviors such that any discrimination guarantees for the affine abstraction also hold for the original nonlinear system. Then, we propose a novel solution in the form of a mixed-integer linear program (MILP) to the active model discrimination problem for uncertain affine models, which includes the affine abstraction and thus, the nonlinear models. Finally, we demonstrate the effectiveness of our approach for identifying the intention of other vehicles in a highway lane changing scenario.

Keywords: Input design, model discrimination, nonlinear systems, affine abstraction

1. INTRODUCTION

Recently, there is much public interest in the integration of smart systems into everyday lives. These systems that include smart homes, smart grids, intelligent transportation and smart cities, are essentially complex, integrated and interconnected engineered systems with multiple operating modes that are often not directly observed or measured; thus, they can be modeled as hidden mode hybrid systems. For example, autonomous vehicles/robots have no access to the intentions or decisions of other vehicles/humans [Sadigh et al. (2016); Yong et al. (2014); Ding et al. (2018)], while smart infrastructures are prone to different fault types [Harirchi and Ozay (2015); Cheong and Manchester (2015)] or attack modes [Pasqualetti et al. (2013); Yong et al. (2015); Harirchi et al. (2017b)]. In these scenarios, approaches for discriminating among these operating modes (or more generally, models of system behaviors) based on noisy observed measurements can have a significant impact on a broad range of applications in robotics, process control, medical devices, fault detection, etc.

Literature Review: The problem of discriminating among a set of models appears in a plethora of research areas such as fault detection, input-distinguishability and mode discernibility of hybrid systems, where the approaches in the literature can be grouped into passive and active methods. Passive discrimination techniques seek the separation of the models regardless of the input [Lou and Si (2009); Rosa and Silvestre (2011); Yong et al. (2014); Harirchi et al. (2017a)], while active methods design a separating input such that the behaviors of different models are distinct. Specifically in the area of input design for active model discrimination, many approaches have been proposed with the goal of finding a small excitation that has a minimal effect on the desired behavior of the system, while guaranteeing the isolation of different fault models [Cheong and Manchester (2015); Šimandl and Punčochář (2009); Nikoukhah and Campbell (2006); Scott et al. (2014); Harirchi et al. (2017b); Ding et al. (2018)]. However, these methods are only applicable for known linear or affine models, and not for nonlinear or uncertain affine models that we consider in this work.

Another set of relevant literature pertains to abstractions of nonlinear systems as linear or affine models that over-approximates all possible original system behaviors, which is a common systematic approximation approach in the literature on hybridization [Asarin et al. (2007); Althoff et al. (2008); Dang et al. (2010)]. This is typically achieved by linear interpolation and over-approximating the interpolation errors as an additive bounded noise term, which may at times be a crude approximation.

Contributions: We propose a novel two-step approach to active model discrimination among a set of uncertain nonlinear models, consisting of the affine abstraction/over-

⋆ This work was supported by Toyota Research Institute ("TRI"). TRI provided funds to assist the authors with their research but this article solely reflects the opinions and conclusions of its authors and not TRI or any other Toyota entity.

approximation and the corresponding input design problem. To the best of our knowledge, this problem is relatively unexplored in the literature.

First, to address the potential conservativeness of conventional abstraction/over-approximation methods, we propose an optimization-based approach to over-approximate nonlinear dynamics by uncertain affine models that are compactly described using interval-valued matrices/vectors, in contrast to only having a interval-valued affine vector. In particular, this uncertain affine model must preserve all the system behaviors of the original nonlinear dynamics such that any model discrimination guarantees for uncertain affine abstraction also hold for the original nonlinear models.

Next, using the resulting set of uncertain affine models, we propose a novel solution to the active model discrimination problem for uncertain affine models, which includes the affine abstraction. We show that this problem can be casted as a mixed-integer linear program (MILP), for which off-the-shelf optimization tools are readily available. To our knowledge, this input design problem for uncertain affine models (an important problem on its own) has not been considered in the literature. Finally, we demonstrate the effectiveness of our approach for identifying the intention of other vehicles in a lane changing scenario.

2. PRELIMINARIES

2.1 Notation and Definitions

Let $x \in \mathbb{R}^n$ denote a vector and $M \in \mathbb{R}^{n\times m}$ a matrix, with transpose $M^\intercal$ and $M \geq 0$ denotes element-wise non-negativity. The vector norm of x is denoted by $\|x\|_i$ with $i \in \{1, 2, \infty\}$, while $\mathbf{0}$, $\mathbb{1}$ and $\mathbb{I}$ represent the vector of zeros, the vector of ones and the identity matrix of appropriate dimensions. The diag and vec operators are defined for a collection of matrices $M_i, i = 1, \ldots, n$ and matrix M as:

$$\mathrm{diag}_{i=1}^n\{M_i\} = \begin{bmatrix} M_1 & & \\ & \ddots & \\ & & M_n \end{bmatrix}, \ \mathrm{vec}_{i=1}^n\{M_i\} = \begin{bmatrix} M_1 \\ \vdots \\ M_n \end{bmatrix},$$
$$\mathrm{diag}_{i,j}\{M_k\} = \begin{bmatrix} M_i & \mathbf{0} \\ \mathbf{0} & M_j \end{bmatrix}, \qquad \mathrm{vec}_{i,j}\{M_k\} = \begin{bmatrix} M_i \\ M_j \end{bmatrix},$$
$$\mathrm{diag}_N\{M\} = \mathbb{I}_N \otimes M, \qquad \mathrm{vec}_N\{M\} = \mathbb{1}_N \otimes M,$$

where $\otimes$ is the Kronecker product. The set of positive integers up to n is denoted by $\mathbb{Z}_n^+$, and the set of non-negative integers up to n is denoted by $\mathbb{Z}_n^0$. Note also the definition of Special Ordered Set of degree 1 (SOS-1) constraints in [Gurobi Optimization (2015)].

2.2 Nonlinear Modeling Framework

Consider N discrete-time affine time-invariant models $\mathcal{G}_i^n$, each with states $\vec{\boldsymbol{x}}_i \in \mathbb{R}^n$, outputs $z_i \in \mathbb{R}^p$, inputs $\vec{\boldsymbol{u}}_i \in \mathbb{R}^m$, process noise $w_i \in \mathbb{R}^{m_w}$ and measurement noise $v_i \in \mathbb{R}^{m_v}$. The models evolve according to the following autonomous nonlinear state equation:

$$\vec{\boldsymbol{x}}_i^+ = \mathfrak{f}_i(\vec{\boldsymbol{x}}_i, \vec{\boldsymbol{u}}_i, w_i), \tag{1}$$

where $\mathfrak{f}_i$ is differentiable. For discrete-time systems, $\vec{\boldsymbol{x}}^+$ denotes the state at the next time instant while for continuous-time systems, $\vec{\boldsymbol{x}}^+ = \dot{\vec{\boldsymbol{x}}}$ is the time derivative of the state. Moreover, the output equation is:

$$z_i = C_i\vec{\boldsymbol{x}}_i + D_i\vec{\boldsymbol{u}}_i + D_{v,i}v_i + g_i. \tag{2}$$

The initial condition for model i, denoted by $\vec{\boldsymbol{x}}_{i,0} = \boldsymbol{x}_i(0)$, is constrained to a polyhedral set with c_0 inequalities:

$$\vec{\boldsymbol{x}}_{i,0} \in \mathcal{X}_0 = \{\vec{\boldsymbol{x}} \in \mathbb{R}^n : P_0\vec{\boldsymbol{x}} \leq p_0\},\ \forall i \in \mathbb{Z}_N^+. \tag{3}$$

The first m_u components of $\vec{\boldsymbol{u}}_i$ are controlled inputs, denoted as $u \in \mathbb{R}^{m_u}$, which are equal for all $\vec{\boldsymbol{u}}_i$, while the other $m_d = m - m_u$ components of $\vec{\boldsymbol{u}}_i$, denoted as $d_i \in \mathbb{R}^{m_d}$, are uncontrolled inputs that are model-dependent. Further, the states $\vec{\boldsymbol{x}}_i$ are partitioned into $x_i \in \mathbb{R}^{n_x}$ and $y_i \in \mathbb{R}^{n_y}$, where $n_y = n - n_x$, as follows:

$$\vec{\boldsymbol{u}}_i = \begin{bmatrix} u \\ d_i \end{bmatrix}, \vec{\boldsymbol{x}}_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}. \tag{4}$$

The states x_i and y_i represent subsets of the states $\vec{\boldsymbol{x}}_i$ that are the 'responsibilities' of the controlled and uncontrolled inputs, u and d_i, respectively. The term 'responsibility' in this paper is to be interpreted as u and d_i, respectively, having to independently satisfy the following polyhedral state constraints with c_x and c_y inequalities:

$$x_i \in \mathcal{X}_{x,i} = \{x \in \mathbb{R}^{n_x} : P_{x,i}x \leq p_{x,i}\}, \tag{5}$$
$$y_i \in \mathcal{X}_{y,i} = \{y \in \mathbb{R}^{n_y} : P_{y,i}y \leq p_{y,i}\}, \tag{6}$$

subject to constrained inputs described by polyhedral sets with c_u and c_d inequalities, respectively:

$$u \in \mathcal{U} = \{u \in \mathbb{R}^{m_u} : Q_u u \leq q_u\}, \tag{7}$$
$$d_i \in \mathcal{D}_i = \{d \in \mathbb{R}^{m_{d_i}} : Q_{d,i}d \leq q_{d,i}\}. \tag{8}$$

Further, $\boldsymbol{\mathcal{X}_i} \triangleq \mathcal{X}_{x,i} \times \mathcal{X}_{y,i}$ and $\boldsymbol{\mathcal{U}_i} \triangleq \mathcal{U} \times \mathcal{D}_i$. On the other hand, the process noise w_i and measurement noise v_i are also polyhedrally constrained with c_w and c_y inequalities:

$$w_i \in \mathcal{W}_i = \{w \in \mathbb{R}^{m_w} : Q_{w,i}w \leq q_{w,i}\}, \tag{9}$$
$$v_i \in \mathcal{V}_i = \{v \in \mathbb{R}^{m_v} : Q_{v,i}v \leq q_{v,i}\}, \tag{10}$$

and have no responsibility to satisfy any state constraints. The readers are referred to [Remark 1, Ding et al. (2018)] for a description of the well-posedness of the formulation.

2.3 Uncertain Affine Modeling Framework

We also consider the uncertain affine modeling framework, that we will show in Section 4 to be a good representation for abstracting/over-approximating nonlinear models. Specifically, consider N discrete-time uncertain affine models $\mathcal{G}_i^a = (\mathcal{A}_i, \mathcal{B}_i, \mathcal{B}_{w,i}, C_i, D_i, D_{v,i}, \mathcal{F}_i, g_i)$, each with states $\vec{\boldsymbol{x}}_i \in \mathbb{R}^n$, outputs $z_i \in \mathbb{R}^p$, inputs $\vec{\boldsymbol{u}}_i \in \mathbb{R}^m$, process noise $w_i \in \mathbb{R}^{m_w}$ and measurement noise $v_i \in \mathbb{R}^{m_v}$. The models evolve according to the state and output equations:

$$\vec{\boldsymbol{x}}_i(k+1) \in \mathcal{C}_i(k), \tag{11}$$
$$z_i(k) = C_i\vec{\boldsymbol{x}}_i(k) + D_i\vec{\boldsymbol{u}}_i(k) + D_{v,i}v_i(k) + g_i, \tag{12}$$

where $\mathcal{C}_i(k)$ is the *cover* of $\underline{\boldsymbol{x}}_i(k+1)$ in given domains $\vec{\boldsymbol{x}}_i \in \boldsymbol{\mathcal{X}}$, $\vec{\boldsymbol{u}}_i \in \boldsymbol{\mathcal{U}}$, $w_i \in \mathcal{W}_i$, defined by

$$\begin{aligned} &\mathcal{C}_i(k) \triangleq \{\vec{\boldsymbol{x}}_i(k+1) \in \boldsymbol{\mathcal{X}} | \\ &\quad \underline{A}_i\vec{\boldsymbol{x}}_i(k) + \underline{B}_{u,i}\vec{\boldsymbol{u}}_i(k) + \underline{B}_{d,i}d_i(k) + \underline{B}_{w,i}w_i(k) + \underline{f}_i \leq \vec{\boldsymbol{x}}_i(k+1) \\ &\quad \leq \overline{A}_i\vec{\boldsymbol{x}}_i(k) + \overline{B}_{u,i}\vec{\boldsymbol{u}}_i(k) + \overline{B}_{d,i}d_i(k) + \overline{B}_{w,i}w_i(k) + \overline{f}_i, \\ &\quad \forall \vec{\boldsymbol{x}}_i \in \boldsymbol{\mathcal{X}}, \underline{\boldsymbol{u}}_i \in \boldsymbol{\mathcal{U}}, w_i \in \mathcal{W}_i\}. \end{aligned} \tag{13}$$

For compactness, we will represent the above model with interval matrices $\mathcal{A}_i = [\underline{A}_i, \overline{A}_i]$, $\mathcal{B}_i = [\underline{B}_i, \overline{B}_i]$, $\mathcal{B}_{w,i} = [\underline{B}_{w,i}, \overline{B}_{w,i}]$ and an interval vector $\mathcal{F}_i = [\underline{f}_i, \overline{f}_i]$. This uncertain model can be obtained by construction, as described in Section 4, and we will use (13) directly for active model discrimination in Section 5.

Based on the partitions of inputs in (4), the corresponding partitioning of the matrices $\underline{B}_i$, $\overline{B}_i$ and D_i are:

$$\underline{B}_i = \begin{bmatrix} \underline{B}_{u,i} & \underline{B}_{d,i} \end{bmatrix}, \overline{B}_i = \begin{bmatrix} \overline{B}_{u,i} & \overline{B}_{d,i} \end{bmatrix}, D_i = \begin{bmatrix} D_{u,i} & D_{d,i} \end{bmatrix}.$$

The states x_i and y_i represent the subset of the states $\vec{\boldsymbol{x}}_i$ that are the 'responsibilities' of the controlled and uncontrolled inputs, u and d_i, respectively. The term 'responsibility' in this paper is to be interpreted as u and d_i, respectively, having to independently satisfy the following polyhedral state constraints (for $k \in \mathbb{Z}_T^+$) with

c_x and c_y inequalities, as is given in (5) and (6), subject to constrained inputs described by polyhedral sets (for $k \in \mathbb{Z}_{T-1}^0$) with c_u and c_d inequalities, respectively, given by (7) and (8). Similarly, the process noise w_i and measurement noise v_i are also polyhedrally constrained with c_w and c_y inequalities, as given in (9) and (10).

In addition, we will consider a time horizon of length T and introduce some time-concatenated notations. The time-concatenated states and outputs are defined as

$$\vec{\boldsymbol{x}}_{i,T} = \mathrm{vec}_{k=0}^{T}\{\vec{\boldsymbol{x}}_i(k)\}, \quad z_{i,T} = \mathrm{vec}_{k=0}^{T}\{z_i(k)\},$$

while the time-concatenated inputs and noises are

$$\vec{\boldsymbol{u}}_{i,T} = \mathrm{vec}_{k=0}^{T-1}\{\vec{\boldsymbol{u}}_i(k)\}, u_T = \mathrm{vec}_{k=0}^{T-1}\{u(k)\}, d_{i,T} = \mathrm{vec}_{k=0}^{T-1}\{d_i(k)\},$$
$$w_{i,T} = \mathrm{vec}_{k=0}^{T-1}\{w_i(k)\}, v_{i,T} = \mathrm{vec}_{k=0}^{T}\{v_i(k)\}.$$

Using the above time-concatenated inputs, noise, states and outputs, the corresponding time-concatenated state and output equations can be written as:

$$\begin{aligned} M_i \vec{\boldsymbol{x}}_{i,T} + \Gamma_{u,i} u_T + \Gamma_{d,i} d_{i,T} + \Gamma_{w,i} w_{i,T} + F_i \leq 0, \\ z_{i,T} = E_i \vec{\boldsymbol{x}}_{i,T} + F_{u,i} u_T + F_{d,i} d_{i,T} + F_{v,i} v_{i,T} + G_i, \end{aligned} \tag{14}$$

where the matrices and vectors M_i, $\Gamma_{u,i}$, $\Gamma_{d,i}$, $\Gamma_{w,i}$, $F_{i,T}$, E_i, $F_{u,i}$, $F_{d,i}$, $F_{d,i}$ and G_i are defined in the appendix.

Given N discrete-time affine models, there are $\mathcal{I} = \binom{N}{2}$ model pairs and let the mode $\iota \in \{1, \cdots, \mathcal{I}\}$ denote the pair of models (i, j). Then, concatenating $\vec{\boldsymbol{x}}_{i,T}$, $x_{i,T}$, $y_{i,T}$, $d_{i,T}$, $z_{i,T}$, $w_{i,T}$ and $v_{i,T}$ for each model pair, we define

$$\vec{\boldsymbol{x}}_T^\iota = \mathrm{vec}_{i,j}\{\vec{\boldsymbol{x}}_{i,T}\}, \; z_T^\iota = \mathrm{vec}_{i,j}\{z_{i,T}\}, \; d_T^\iota = \mathrm{vec}_{i,j}\{d_{i,T}\},$$
$$w_T^\iota = \mathrm{vec}_{i,j}\{w_{i,T}\}, \; v_T^\iota = \mathrm{vec}_{i,j}\{v_{i,T}\}.$$

The states and outputs over the entire time horizon for each mode ι can be written as simple functions of the state $\vec{\boldsymbol{x}}^\iota$, input vectors u_T, d_T^ι, and noise w_T^ι, v_T^ι:

$$\begin{aligned} M^\iota \vec{\boldsymbol{x}}_T^\iota + \Gamma_u^\iota u_T + \Gamma_d^\iota d_T^\iota + \Gamma_w^\iota w_T^\iota + F_T^\iota \leq 0, \\ z_T^\iota = E^\iota \vec{\boldsymbol{x}}_T^\iota + F_u^\iota u_T^\iota + F_d^\iota d_T^\iota + F_v^\iota v_T^\iota + G^\iota. \end{aligned} \tag{15}$$

The matrices and vectors M^ι, Γ_u^ι, Γ_d^ι, Γ_w^ι, F_T^ι, E^ι, F_u^ι, F_d^ι, F_v^ι and G^ι are defined in the appendix. Moreover, the uncertain variables for each mode ι are concatenated as $\overline{x}^\iota = [\vec{\boldsymbol{x}}_T^{\iota\top} \; d_T^{\iota\top} \; w_T^{\iota\top} \; v_T^{\iota\top}]^\top$. We then concatenate the polyhedral state constraints in (5) and (6). First, let

$$\overline{P}_{\underline{x}}^\iota = \mathrm{diag}_{i,j}\,\mathrm{diag}_T\{P_{\underline{x},i}\}, \quad \overline{p}_{\underline{x}}^\iota = \mathrm{vec}_{i,j}\,\mathrm{vec}_T\{p_{\underline{x},i}\},$$

where $P_{\underline{x},i} = \begin{bmatrix} P_{x,i} & 0 \\ 0 & P_{y,i} \end{bmatrix}$ and $p_{\underline{x},i} = \begin{bmatrix} p_{x,i} \\ p_{y,i} \end{bmatrix}$. Further, we concatenate the initial state constraint in (3):

$$\overline{P}_0^\iota = \mathrm{diag}_2\{[P_0 \; 0]\}, \quad \overline{p}_0^\iota = \mathrm{vec}_2\{p_0\}.$$

Similarly, let

$$\overline{Q}_u = \mathrm{diag}_T\{Q_u\}, \; \overline{Q}_\dagger^\iota = \mathrm{diag}_{i,j}\,\mathrm{diag}_T\{Q_{\dagger,i}\},$$
$$\overline{q}_u = \mathrm{vec}_T\{q_u\}, \; \overline{q}_\dagger^\iota = \mathrm{vec}_{i,j}\,\mathrm{vec}_T\{q_{\dagger,i}\}, \; \dagger \in \{d, w, v\}.$$

Then, the input constraints in (7) and (9) for all k are: $\overline{Q}_u u_T \leq \overline{q}_u$ and $\overline{Q}_\dagger^\iota \dagger_T^\iota \leq \overline{q}_\dagger^\iota$. Hence, in terms of $\overline{x}^\iota$, we have a polyhedral constraint of the form $H_{\bar{x}}^\iota \bar{x}^\iota \leq h_{\bar{x}}^\iota$, with

$$H_{\bar{x}}^\iota = \begin{bmatrix} \begin{bmatrix} \bar{P}_0^\iota & 0 \end{bmatrix} \\ \mathrm{diag}\{\bar{P}_x^\iota, \bar{Q}_d^\iota, \bar{Q}_w^\iota, \bar{Q}_v^\iota\} \end{bmatrix}, h_{\bar{x}}^\iota = \begin{bmatrix} \bar{p}_0^{\iota,\top} & \bar{p}_x^{\iota,\top} & \bar{q}_d^{\iota,\top} & \bar{q}_w^{\iota,\top} & \bar{q}_v^{\iota,\top} \end{bmatrix}^\top.$$

Note that the above definitions of matrices and vectors for the *uncertain affine model* are different from the ones in Ding et al. (2018) that were defined for *affine models*.

3. PROBLEM FORMULATION

We consider the active model discrimination problem among a finite number of uncertain nonlinear models, i.e., to design a *separating* admissible input for the system such that when the system is excited with this input, any observed trajectory is consistent with only one model, regardless of any realization of uncertain parameters. In addition, we minimize a cost function $J(u)$, as follows:

Problem 1. (Active Model Discrimination). Given N well-posed nonlinear models $\mathcal{G}_i$, and state, input and noise constraints, (3),(6),(8)-(10), find an optimal input sequence u_T^* to minimize a given cost function $J(u_T)$ such that for all possible initial states $\vec{\boldsymbol{x}}_0$, uncontrolled input d_T, process noise w_T and measurement noise v_T, only one model is valid, i.e., the output trajectories of any pair of models have to differ in at least one time instance.

The above problem is in general computationally intractable with no optimality guarantees, as it would result in a mixed-integer nonlinear program (MINLP). Thus, we propose a two-step approach to tackle this problem. The first subproblem (Problem 1.1) is to over-approximate the nonlinear dynamics in (1) by uncertain affine models (for each $i \in \mathbb{Z}_{\mathrm{N}}^+$; subscript i omitted below for brevity), as abstractions that preserve (or "contain") all its system behaviors. This is a common systematic approximation approach in the literature on hybridization [Asarin et al. (2007); Althoff et al. (2008); Dang et al. (2010)].

Problem 1.1. (Affine Abstraction). Given a *nonlinear* n-dimensional vector field $\mathfrak{f}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{u}}, w)$ with (polytopic) domain $\vec{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}$, $\vec{\boldsymbol{u}} \in \boldsymbol{\mathcal{U}}$, $w \in \mathcal{W}$, find two n dimensional affine planes (i.e., upper and lower planes defined by matrices $\overline{A}, \underline{A}, \overline{B}, \underline{B}, \overline{B}_w, \underline{B}_w$ and vectors $\overline{f}$, $\underline{f}$) such that they "contain" (i.e., upper- and lower bound) the given vector field with minimum separation, as expressed by the following:

$$\min_{\overline{A},\underline{A},\overline{B},\underline{B},\overline{B}_w,\underline{B}_w,\overline{f},\underline{f}\widetilde{A},\widetilde{B},\widetilde{B}_w,\widetilde{f}} \widetilde{A} + \lambda_1 \widetilde{B} + \lambda_2 \widetilde{B}_w + \lambda_3 \widetilde{f}$$

subject to

$$\begin{aligned} \underline{A}\,\vec{\boldsymbol{x}} + \underline{B}\,\vec{\boldsymbol{u}} + \underline{B}_w w + \underline{f} \leq \mathfrak{f}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{u}}, w), \\ \mathfrak{f}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{u}}, w) \leq \overline{A}\,\vec{\boldsymbol{x}} + \overline{B}\,\vec{\boldsymbol{u}} + \overline{B}_w w + \overline{f}, \end{aligned} \forall (\vec{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}, \vec{\boldsymbol{u}} \in \boldsymbol{\mathcal{U}}, \; w \in \mathcal{W}), \tag{16a}$$

$$\begin{aligned} \|\overline{A} - \underline{A}\| \leq \widetilde{A}, \|\overline{B} - \underline{B}\| \leq \widetilde{B}, \|\overline{B}_w - \underline{B}_w\| \leq \widetilde{B}_w, \\ \|\overline{f} - \underline{f}\| \leq \widetilde{f}, \overline{A} \geq \underline{A}, \overline{B} \geq \underline{B}, \overline{B}_w \geq \underline{B}_w, \overline{f} \geq \underline{f}, \end{aligned} \tag{16b}$$

where λ_1, λ_2 and λ_3 are tuning weights and suitable norms are chosen based on the application at hand.

Then, using the resulting set of uncertain affine models that conform to the modeling framework in Section 2.3, we consider the following second subproblem (Problem 1.2):

Problem 1.2. (Active Affine Model Discrimination). Given N well-posed affine models $\mathcal{G}_i^a$, and state, input and noise constraints, (3),(6),(8)-(10), find an optimal input sequence u_T^* to minimize a given cost function $J(u_T)$ such that for all possible initial states $\vec{\boldsymbol{x}}_0$, uncontrolled input d_T, process noise w_T and measurement noise v_T, only one model is valid, i.e., the output trajectories of any pair of models have to differ in at least one time instance. The optimization problem can be formally stated as follows:

$$\min_{u_T, x_T} J(u_T) \text{ s.t. } \quad \forall k \in \mathbb{Z}_{T-1}^0 : (7) \text{ hold}, \tag{17a}$$

$$\left.\begin{aligned} \forall i, j \in \mathbb{Z}_{\mathrm{N}}^+, i < j, \forall k \in \mathbb{Z}_T^0, \\ \forall \vec{\boldsymbol{x}}_T, y_T, d_T, w_T, v_T : \\ (3),(5),(6),(8)\text{-}(10),(15) \text{ hold} \end{aligned}\right\} : \begin{aligned} \exists k \in \mathbb{Z}_T^0, \\ z_i(k) \neq z_j(k). \end{aligned} \tag{17b}$$

Using the two-step design, the discrimination guarantees in Problem 1.2 also hold for the original nonlinear models because the affine abstraction in Problem 1.1 is designed such that any properties that hold for the abstraction also hold for the original nonlinear models. Our approach has the advantage of "divorcing" the difficulty associated with optimization problems with nonlinear constraints from the active model discrimination problem, providing insight into whether the nonlinearity is the bottleneck or if the active discrimination problem itself is ill-posed. However, the solution is, by construction, suboptimal. Nonetheless, the designed separating input guarantees that all output trajectories of all the nonlinear models are distinguishable from each other under any realization of uncertainties in the initial condition, model discrepancies or noise.

4. AFFINE ABSTRACTION

In this section, we address Problem 1.1. Since the robust formulation given in Problem 1.1 cannot be directly implemented using standard optimization packages, we will convert the problem into a more amenable form such that off-the-shelf optimization tools can be applied. For simplicity, we first describe our approach for a 1-dimensional vector field with interval domains, before describing how this can be extended to higher dimensional systems. Moreover, we will assume in the following discussion that there is at most one local optimum in the domain $\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{U}}, \mathcal{W}$. This can also be extended to the case with multiple local optima.

Theorem 1. (Affine Abstraction). Given a *nonlinear* 1-dimensional differentiable vector field $\mathfrak{f}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{u}}, w)$ with interval domains $\vec{\boldsymbol{x}} \in \boldsymbol{\mathcal{X}}$, $\vec{\boldsymbol{u}} \in \boldsymbol{\mathcal{U}}$, $w \in \mathcal{W}$, two affine planes (i.e., upper and lower planes defined by $\overline{A}, \underline{A}, \overline{B}, \underline{B}, \overline{B}_w, \underline{B}_w$ and vectors $\overline{f}$, $\underline{f}$) that contain the given vector field with minimum separation are solutions to:

$$\min_{\substack{\overline{A}, \underline{A}, \overline{B}, \underline{B}, \overline{B}_w, \underline{B}_w, \overline{f}, \underline{f} \widetilde{A}, \widetilde{B}, \widetilde{B}_w, \widetilde{f}, \\ \vec{\boldsymbol{x}}_u, \vec{\boldsymbol{u}}_u, w_u, \vec{\boldsymbol{x}}_b, \vec{\boldsymbol{u}}_b, w_b}} \widetilde{A} + \lambda_1 \widetilde{B} + \lambda_2 \widetilde{B}_w + \lambda_3 \widetilde{f}$$

subject to

$$\begin{aligned} &\underline{A}\,\vec{\boldsymbol{x}} + \underline{B}\,\vec{\boldsymbol{u}} + \underline{B}_w w + \underline{f} \leq \mathfrak{f}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{u}}, w), \\ &\mathfrak{f}(\vec{\boldsymbol{x}}, \vec{\boldsymbol{u}}, w) \leq \overline{A}\,\vec{\boldsymbol{x}} + \overline{B}\,\vec{\boldsymbol{u}} + \overline{B}_w w + \overline{f}, \end{aligned} \quad \forall (\vec{\boldsymbol{x}} \in \widetilde{\boldsymbol{\mathcal{X}}}, \vec{\boldsymbol{u}} \in \widetilde{\boldsymbol{\mathcal{U}}}, w \in \widetilde{\mathcal{W}}), \tag{18a}$$

$$\begin{gathered} \underline{A} - \nabla_{\vec{\boldsymbol{x}}}\, \mathfrak{f}(\vec{\boldsymbol{x}}_b, \vec{\boldsymbol{u}}_b, w_b) = 0, \ \underline{B} - \nabla_{\vec{\boldsymbol{u}}}\, \mathfrak{f}(\vec{\boldsymbol{x}}_b, \vec{\boldsymbol{u}}_b, w_b) = 0, \\ \underline{B}_w - \nabla_w\, \mathfrak{f}(\vec{\boldsymbol{x}}_b, \vec{\boldsymbol{u}}_b, w_b) = 0, \\ ((\vec{\boldsymbol{x}}_b \in \boldsymbol{\mathcal{X}} \wedge \vec{\boldsymbol{u}}_b \in \boldsymbol{\mathcal{U}} \wedge w_b \in \mathcal{W} \wedge (\underline{A}\,\vec{\boldsymbol{x}}_b + \underline{B}\,\vec{\boldsymbol{u}}_b \\ + \underline{B}_w w_b + \underline{f} \leq \mathfrak{f}(\vec{\boldsymbol{x}}_b, \vec{\boldsymbol{u}}_b, w_b))) \vee \vec{\boldsymbol{x}}_b \notin \boldsymbol{\mathcal{X}} \vee \vec{\boldsymbol{u}}_b \notin \boldsymbol{\mathcal{U}} \vee w_b \notin \mathcal{W}), \end{gathered} \tag{18b}$$

$$\begin{gathered} \overline{A} - \nabla_{\vec{\boldsymbol{x}}}\, \mathfrak{f}(\vec{\boldsymbol{x}}_u, \vec{\boldsymbol{u}}_u, w_u) = 0, \ \overline{B} - \nabla_{\vec{\boldsymbol{u}}}\, \mathfrak{f}(\vec{\boldsymbol{x}}_u, \vec{\boldsymbol{u}}_u, w_u) = 0, \\ \overline{B}_w - \nabla_w\, \mathfrak{f}(\vec{\boldsymbol{x}}_u, \vec{\boldsymbol{u}}_u, w_u) = 0, \\ ((\vec{\boldsymbol{x}}_u \in \boldsymbol{\mathcal{X}} \wedge \vec{\boldsymbol{u}}_u \in \boldsymbol{\mathcal{U}} \wedge w_u \in \mathcal{W} \wedge (\underline{A}\,\vec{\boldsymbol{x}}_u + \underline{B}\,\vec{\boldsymbol{u}}_u \\ + \underline{B}_w w_u + \underline{f} \geq \mathfrak{f}(\vec{\boldsymbol{x}}_u, \vec{\boldsymbol{u}}_u, w_u))) \vee \vec{\boldsymbol{x}}_u \notin \boldsymbol{\mathcal{X}} \vee \vec{\boldsymbol{u}}_u \notin \boldsymbol{\mathcal{U}} \vee w_u \notin \mathcal{W}), \end{gathered} \tag{18c}$$

$$\begin{aligned} &\|\overline{A} - \underline{A}\| \leq \widetilde{A}, \|\overline{B} - \underline{B}\| \leq \widetilde{B}, \|\overline{B}_w - \underline{B}_w\| \leq \widetilde{B}_w, \\ &\|\overline{f} - \underline{f}\| \leq \widetilde{f}, \overline{A} \geq \underline{A}, \overline{B} \geq \underline{B}, \overline{B}_w \geq \underline{B}_w, \overline{f} \geq \underline{f}, \end{aligned} \tag{18d}$$

where λ_1, λ_2, λ_3 are tuning weights (chosen based on the application), $\widetilde{\boldsymbol{\mathcal{X}}}$, $\widetilde{\boldsymbol{\mathcal{U}}}$, $\widetilde{\mathcal{W}}$ are the sets of endpoints of the interval domains $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{U}}$, $\mathcal{W}$, respectively, $\vec{\boldsymbol{x}}_u$, $\vec{\boldsymbol{u}}_u$ and w_u are the local optimum for the difference between the upper plane and function $\mathfrak{f}$, and similarly, $\vec{\boldsymbol{x}}_b$, $\vec{\boldsymbol{u}}_b$ and w_b for the bottom plane, while $\wedge$ and $\vee$ are logical `AND` and `OR` operators [1] and ∇_* denotes the derivative operator with respect to variable $i \in \{\vec{\boldsymbol{x}}, \vec{\boldsymbol{u}}, w\}$.

Proof. The optimization formulation above is as in Problem 1.1 except that the semi-infinite constraints (16a) are replaced by readily implementable constraints (18a), (18c), (18b). This is possible because the maximum of a differentiable (thus, continuous) function over a closed domain is the maximum of its local optima (i.e., minima, maxima or saddle points) in its interior (i.e., (18b)) and the maximum over its boundaries (i.e., its endpoints; (18a)). The same holds for minimization over a closed domain, resulting in (18a) and (18c) for the lower plane.

Eq. (18a) is required such that the upper and lower planes upper- and lower-bound each boundary/endpoint of the domains. On the other hand, Eqs. (18b) and (18c) are such that the local optima, which are given by their first order necessary condition (i.e., their first order derivative is set to zero), are also upper- and lower-bounded by the two planes, respectively. The constraints with logical operators are to be understood to be in conjunction with the other constraints, and has the interpretation that only the local optima in the given domain needs to be upper- or lower-bounded by the two planes. Finally, Eq. (18d) along with the objective function ensures that the upper plane remains above the lower plane and that the separation between the two planes is as small as possible. □

To extend the above formulation to higher dimensions, the key difference from the 1-dimensional case is that the boundaries of the domain are higher dimensional facets as opposed to line segments. To deal with this, the above procedure of replacing the maximization or minimization over a domain with the maximizing or minimizing over the local optima (using first order optimality condition for constrained optimization) and its boundaries as in (18a), (18c), (18b) can be repeated to recursively reduce the dimension of the boundaries by one until we obtain line segments. This procedure may be tedious to implement for systems with high dimensions and domains with many facets, hence, in practice, we may start with only constraints on the vertices and iteratively add facets of the domain of interest when the resulting planes are found to intersect them. Further effort to reduce the complexity of this procedure is the subject of ongoing research.

5. ACTIVE MODEL DISCRIMINATION APPROACH

We now extend and modify the approach in Ding et al. (2018) to solve Problem 1.2 with given uncertain affine models from Section 4. This approach relies on formulating the problem as a bi-level optimization problem which is then converted to a single level MILP using KKT conditions, for which off-the-shelf softwares are readily available, e.g., [Gurobi Optimization (2015)]. For the sake of clarity, we will defer the definitions of certain matrices in the following results to the appendix. Moreover, for brevity, the proofs of this approach are omitted, as they follow similar steps to the proofs in [Ding et al. (2018)].

Theorem 2. (Discriminating Input Design as an MILP). Given a separability index ϵ, the active model discrimination problem (Problem 1.2) is equivalent (up to ϵ) to:

$$\min_{u_T, \delta^\iota,, \bar{x}^\iota, \mu_1^\iota,, \mu_2^\iota, \mu_3^\iota} J(u_T) \tag{P_{DID}}$$

$$\begin{aligned} \text{s.t.} \quad & \bar{Q}_u u_T \leq \bar{q}_u, \\ \forall \iota \in \mathbb{Z}_{\mathcal{I}}^+ : \quad & \delta^\iota(u_T) \geq \epsilon, \quad 0 = 1 - \mu_3^{\iota\,\mathrm{T}} \mathbb{1}, \\ & 0 = \textstyle\sum_{i=1}^{i=\kappa} \mu_{1,i}^\iota H_{\bar{x}}^\iota(i,m) + \sum_{j=1}^{j=\xi} \mu_{2,j}^\iota R_1^\iota(j,m) \\ & \quad + \textstyle\sum_{k=1}^{k=\rho} \mu_{3,k}^\iota R_2^\iota(k,m), \forall m = 1, \cdots, \eta, \\ & \widetilde{H}_{\bar{x},i}^\iota \bar{x}^\iota - h_{\bar{x},i}^\iota \leq 0, \mu_{1,i}^\iota \geq 0, \forall i = 1, \dots \kappa, \\ & \widetilde{R}_{1,j}^\iota \bar{x}^\iota - r_{1,j}^\iota - S_{1,j}^\iota u_T \leq 0, \mu_{2,j}^\iota \geq 0, \forall j = 1, \dots \xi, \\ & \widetilde{R}_{2,k}^\iota \bar{x}^\iota - \delta^\iota - r_{2,k}^\iota - S_{2,k}^\iota u_T \leq 0, \mu_{3,k}^\iota \geq 0, \forall k = 1, \dots \rho, \\ \forall \iota \in \mathbb{Z}_{\mathcal{I}}^+, \forall i \in \mathbb{Z}_\kappa^+ : \quad & \text{SOS-1} : \{\mu_{1,i}^\iota, \widetilde{H}_{\bar{x},i}^\iota \bar{x}^\iota - h_{\bar{x},i}^\iota\}, \\ \forall \iota \in \mathbb{Z}_{\mathcal{I}}^+, \forall j \in \mathbb{Z}_\xi^+ : \quad & \text{SOS-1} : \{\mu_{2,j}^\iota, \widetilde{R}_{1,j}^\iota \bar{x}^\iota - r_{1,j}^\iota - \tilde{S}_{1,j}^\iota u_T\}, \\ \forall \iota \in \mathbb{Z}_{\mathcal{I}}^+, \forall j \in \mathbb{Z}_\rho^\rho : \quad & \text{SOS-1} : \{\mu_{3,k}^\iota, \widetilde{R}_{2,k}^\iota \bar{x}^\iota - \delta^\iota - r_{2,k}^\iota - \tilde{S}_{2,k}^\iota u_T\}, \end{aligned}$$

[1] Logical operators are implementable by off-the-shelf softwares, e.g., YALMIP [Löfberg (2004)], and if needed, converted to mixed-integer constraints.

where $\mu^{\iota}_{1,i}$, $\mu^{\iota}_{2,j}$ and $\mu^{\iota}_{3,j}$ are dual variables, while $\widetilde{H}^{\iota}_{\bar{x},i}$ is the i-th row of $H^{\iota}_{\bar{x}}$, $\widetilde{R}^{\iota}_{1,j}$ and $\widetilde{S}^{\iota}_{1,j}$ are the j-th row of R^{ι}_1 and S^{ι}_1, respectively, $\widetilde{R}^{\iota}_{1,k}$ and $\widetilde{S}^{\iota}_{2,k}$ are the k-th row of R^{ι}_2 and S^{ι}_2, respectively, $\eta = \mathcal{I}T(n + m_d + m_w + m_v)$ is the number of columns of $H^{\iota}_{\bar{x}}$, $\kappa = 2\mathcal{I}T(c_0 + c_d + c_w + c_v)$ is the number of rows of $H^{\iota}_{\bar{x}}$, $\xi = 2\mathcal{I}T(c_x + c_y)$ is the number of rows of R^{ι}_1 and $\rho = 2\mathcal{I}Tp$ is the number of rows of R^{ι}_2.

Note that the result above is an extension of Ding et al. (2018). Although the notations of matrices and vectors look similar, their definitions are rather different.

6. APPLICATION EXAMPLE: ACTIVE INTENTION IDENTIFICATION IN LANE CHANGE SCENARIO

In this section, we apply the active model discrimination approach proposed in the previous sections to design a separating control input that, in conjunction with a modified model invalidation algorithm (e.g., [Harirchi and Ozay (2015); Harirchi et al. (2017a, 2018)]), can be used for active intention identification in a highway lane change scenario (cf. [Figure 1, Ding et al. (2018)]).

6.1 Vehicle and Intention Models

The dynamics of the ego vehicle is of the Dubins car [Dubins (1957)] with acceleration input:

$$\dot{x}_e = v_e \cos\phi_e,\ \dot{y}_e = v_e \sin\phi_e,\ \dot{v}_e = u_1 + w_1,\ \dot{\phi}_e = u_2 + w_2,$$

while the other car dynamics is described by:

$$\dot{x}_o = v_o,\ \dot{v}_o = d_i + w_3,$$

where x_e and y_e are the longitudinal and lateral coordinates of the ego car, v_e is its speed and ϕ_e is its heading angle, while x_o is the longitudinal coordinate of the other car with its speed given by v_o (no lateral movement), with process noise signals $w_j, j \in \{1, 2, 3\}$. $u_1 \in \mathcal{U}_a \triangleq [-7.848, 3.968]$ and $u_2 \in \mathcal{U}_s \triangleq [-0.44, 0.44]$ are the acceleration and steering inputs of the ego car, whereas d_i is the (uncontrolled) acceleration input of the other car for each intention $i \in \{I, C, M\}$, corresponding to an Inattentive, Cautious or Malicious driver.

The **Inattentive** driver is unaware of the ego car and maintains his speed using an acceleration input which lies in a small range $d_I \in \mathcal{D}_I \triangleq [-0.392, 0.198]$. The **Cautious** driver tends to yield the lane to the ego car with the input equal to $d_C \triangleq -K_{d,C}(v_e - v_o) + L_{p,C}\phi_e + L_{d,C}\dot{\phi}_e + \tilde{d}_C$, where $\dot{\phi}_e = u_2 + w_2$, $K_{d,C} = 1$, $L_{p,C} = 12$ and $L_{d,C} = 14$ are PD controller parameters and the input uncertainty is $\tilde{d}_C \in \mathcal{D}_C = \mathcal{D}_I$. Finally, the **Malicious** driver does not want to yield the lane and attempts to cause a collision with input equal to $d_M \triangleq K_{d,M}(v_e - v_o) - L_{p,M}\phi_e - L_{d,M}\dot{\phi}_e + \tilde{d}_M$, if provoked, where $\dot{\phi}_e = u_2 + w_2$, $K_{d,C} = 0.9$, $L_{p,C} = 12$ and $L_{d,C} = 14$ are PD controller parameters and the input uncertainty satisfies $\tilde{d}_M \in \mathcal{D}_M = \mathcal{D}_I$.

Without loss of generality, we assume that the initial position and heading angle of the ego car are 0, while the initial velocities match typical speed limits of the highway. Moreover, both cars are close to the center of their lanes that are 3.2 m wide. Thus, the initial conditions are:

$$v_e(0) \in [23, 27]\tfrac{m}{s}, \quad y_e(0) \in [1.5, 1.7]m,$$
$$v_o(0) \in [23, 27]\tfrac{m}{s}, \quad x_o(0) \in [10, 12]m.$$

Further, the velocity of the ego vehicle is constrained to be between $[20, 30]\frac{m}{s}$ at all times (in order to obey the speed limit of a highway), its heading angle is between $[-0.44, 0.44]rad$ and its lateral position is constrained between $[0.3, 2.5]m$. Process and measurement noise signals are bounded with a range of $[-0.01, 0.01]$ and the separability threshold is set to $\epsilon = 0.5\frac{m}{s}$. Moreover, we assume the extreme scenario where only noisy observation of other car's velocity is observed $z = v_o + v$.

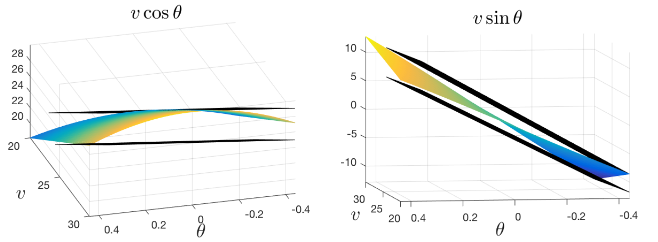

Fig. 1. Affine abstraction/over-approximation of Dubins vehicle dynamics such that the true nonlinear system behavior is contained/enveloped by the abstraction.

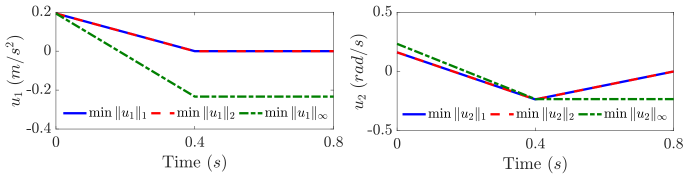

Fig. 2. Effect of different choices of objective functions on the resulting separating inputs. (Note that the control inputs with 1- and 2-norms as objective functions are near identical.)

6.2 Affine Abstraction

The Dubins vehicle and intention models above are nonlinear. Hence, we resort to the approach in Section 4 to obtain affine abstractions of the models. Since the nonlinearity only affects the speed v_e and the heading angle ϕ_e, we first define a suitable domain that is appropriate for the lane changing scenario. Specifically, we consider the speed of the ego car to be between 20 m/s to 30 m/s (72 to 108 km/h) and a heading angle range of $-25°$ to $25°$ ($[-0.44, 0.44]rad$). Using this domain, we can obtain an affine abstraction for the reduced 2-dimensional system with $\lambda_1 = 1000, \lambda_2 = \lambda_3 = 0$ and ∞-norms, as illustrated by Figure 1, and using the compact interval matrix representation, the abstracted open-loop model is given by:

$$\begin{aligned}
\mathcal{A} = [\underline{A}, \overline{A}] \quad &= [(1,3) : [0.9947, 0.9956]; (1,4) : 2.3821;\\
&\quad (2,3) : [-0.1028, -0.0928];\\
&\quad (2,4) : [23.0325, 25.2673]; (5,6) : 1]_{6\times 6},\\
\mathcal{B} = \mathcal{B}_w = B \quad &= [(3,1) : 1; (4,2) : 1; (6,3) : 1]_{6\times 3},\\
\mathcal{F} = [\underline{f}, \overline{f}] \quad &= [(1,1) : [-3.7471, 0.2859];\\
&\quad (1,2) : [0.4413, 4.4754]]_{6\times 1},
\end{aligned}$$

where we used a sparse matrix notation with the size indicted in the subscript.

Combining the abstraction with the intention models and using Euler method for time-discretization with sampling time $\delta t = 0.4$ s, we have the following intention models:

Inattentive Driver $(i = I)$:

$$\underline{A}_I = \mathbb{I} + \delta t\underline{A},\ \overline{A}_I = \mathbb{I} + \delta t\overline{A},$$
$$\underline{B}_I = \overline{B}_I = \underline{B}_{w,I} = \overline{B}_{w,I} = \delta t B,$$
$$C_I = [(1,6) : 1]_{1\times 6},\ D_I = 0,\ D_{v,I} = 1,\ \underline{f}_I = \delta t\underline{f},\ \overline{f}_I = \delta t\overline{f}.$$

Cautious Driver $(i = C)$:

$$\tilde{A}_C = [(6,3) : -K_{d,C}; (6,4) : L_{d,C}; (6,6) : K_{d,C}]_{6\times 6},$$
$$\tilde{B}_C = [(6,2) : L_{p,C}]_{6\times 3},$$
$$\underline{A}_C = \mathbb{I} + \delta t(\underline{A} + \tilde{A}_C),\ \overline{A}_C = \mathbb{I} + \delta t(\overline{A} + \tilde{A}_C),$$
$$\underline{B}_C = \overline{B}_C = \underline{B}_{w,C} = \overline{B}_{w,C} = \delta t(B + \tilde{B}_C),$$
$$C_C = C_I,\ D_C = 0,\ D_{v,C} = 1,\ \underline{f}_C = \delta t\underline{f},\ \overline{f}_C = \delta t\overline{f}.$$

Malicious Driver $(i = M)$:

$$\tilde{A}_M = [(6,3) : -K_{d,M}; (6,4) : L_{d,M}; (6,6) : K_{d,M}]_{6\times 6},$$
$$\tilde{B}_M = [(6,2) : L_{p,M}]_{6\times 3},$$
$$\underline{A}_M = \mathbb{I} + \delta t(\underline{A} - \tilde{A}_M),\ \overline{A}_M = \mathbb{I} + \delta t(\overline{A} - \tilde{A}_M),$$
$$\underline{B}_M = \overline{B}_M = \underline{B}_{w,M} = \overline{B}_{w,M} = \delta t(B - \tilde{B}_M),$$
$$C_M = C_I,\ D_M = 0,\ D_{v,M} = 1,\ \underline{f}_M = \delta t\underline{f},\ \overline{f}_M = \delta t\overline{f}.$$

6.3 Active Nonlinear Model Discrimination

Next, we apply the active model discrimination approach in Section 5 to the above uncertain affine models for the three intentions of the other driver in a highway lane changing scenario. Figure 2 shows the active inputs of the ego car to discern the other car's intention based on its response when using various objective functions. When comparing the solutions with different norms as the objective functions, we observe that $\|u\|_1$ reduces the sum of absolute values, thus keeping all the data points as close to zero as possible, $\|u\|_2$ minimizes energy and may be desirable to reduce fuel consumption, and $\|u\|_\infty$ ensures comfort by minimizing the maximum absolute input values. In all solutions, the ego car accelerates and turns towards the other car's lane, and then either maintains its speed or decelerates while re-aligning to its lane, although now laterally shifted toward the other car's lane. Throughout the maneuver, the inattentive car coasts with small accelerations, the cautious car slows down when it observes the ego car turning to its lane and the malicious car tries to match the ego car's position, causing a collision.

The obtained optimal separating input is then applied to the ego vehicle in real-time. After measurements are recorded for $T = 3$, the passive model discrimination approach based on model invalidation of each intention model (e.g., [Harirchi and Ozay (2015); Harirchi et al. (2018)]) can be applied to identify the intention of the other vehicle. This is guaranteed to find the true intention by design (see above definition of separating input).

7. CONCLUSION

This work considered the novel design of separating input signals in order to discriminate among a finite number of uncertain nonlinear models, using a two-step approach. First, we developed an optimization-based approach that is less conservative than existing methods to over-approximate nonlinear dynamics by uncertain affine models that contain all the system behaviors of the original nonlinear system. Then, we proposed one of the first active model discrimination algorithms for uncertain affine models, which includes the affine abstraction, hence, the nonlinear models. Finally, we demonstrated our approach on an example of intent estimation/identification in a lane changing scenario on a highway. For future work, we are interested to overcome the complexity of our affine abstraction approach and to reduce its optimality gap.

REFERENCES

Althoff, M., Stursberg, O., and Buss, M. (2008). Reachability analysis of nonlinear systems with uncertain parameters using conservative linearization. In *IEEE CDC*, 4042–4048.

Asarin, E., Dang, T., and Girard, A. (2007). Hybridization methods for the analysis of nonlinear systems. *Acta Inf.*, 43(7), 451–476.

Cheong, S. and Manchester, I. (2015). Input design for discrimination between classes of LTI models. *Automatica*, 53, 103–110.

Dang, T., Maler, O., and Testylier, R. (2010). Accurate hybridization of nonlinear systems. In *ACM HSCC*, 11–20.

Ding, Y., Harirchi, F., Yong, S.Z., Jacobsen, E., and Ozay, N. (2018). Optimal input design for affine model discrimination with applications in intention-aware vehicles. In *ACM/IEEE ICCPS*, 5161–5167.

Dubins, L.E. (1957). On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents. *American J. of Math.*, 79, 497–516.

Gurobi Optimization, I. (2015). Gurobi optimizer reference manual.

Harirchi, F. and Ozay, N. (2015). Model invalidation for switched affine systems with applications to fault and anomaly detection. *IFAC-PapersOnLine*, 48(27), 260–266.

Harirchi, F., Yong, S.Z., and Ozay, N. (2017a). Guaranteed fault detection and isolation for switched affine models. In *IEEE CDC*.

Harirchi, F., Yong, S.Z., Jacobsen, E., and Ozay, N. (2017b). Active model discrimination with applications to fraud detection in smart buildings. *IFAC-PapersOnLine*, 50(1), 9527–9534.

Harirchi, F., Yong, S.Z., and Ozay, N. (2018). Passive diagnosis of hidden-mode switched affine models with detection guarantees via model invalidation. In *Diagnosability, Security and Safety of Hybrid Dynamic and Cyber-Physical Systems*, 227–251. Springer.

Löfberg, J. (2004). YALMIP: A toolbox for modeling and optimization in MATLAB. In *CACSD*. Taipei, Taiwan.

Lou, H. and Si, P. (2009). The distinguishability of linear control systems. *Nonlinear Analysis: Hybrid Systems*, 3(1), 21–38.

Nikoukhah, R. and Campbell, S. (2006). Auxiliary signal design for active failure detection in uncertain linear systems with a priori information. *Automatica*, 42(2), 219–228.

Pasqualetti, F., Dörfler, F., and Bullo, F. (2013). Attack detection and identification in cyber-physical systems. *IEEE Trans. on Aut. Cont.*, 58(11), 2715–2729.

Rosa, P. and Silvestre, C. (2011). On the distinguishability of discrete linear time-invariant dynamic systems. In *IEEE CDC-ECC*, 3356–3361.

Sadigh, D., Sastry, S., Seshia, S., and Dragan, A. (2016). Information gathering actions over human internal state. In *IEEE/RSJ IROS*, 66–73.

Scott, J., Findeisen, R., Braatz, R.D., and Raimondo, D. (2014). Input design for guaranteed fault diagnosis using zonotopes. *Automatica*, 50(6), 1580–1589.

Šimandl, M. and Punčochář, I. (2009). Active fault detection and control: Unified formulation and optimal design. *Automatica*, 45(9), 2052–2059.

Yong, S.Z., Zhu, M., and Frazzoli, E. (2014). Generalized innovation and inference algorithms for hidden mode switched linear stochastic systems with unknown inputs. In *IEEE CDC*, 3388–3394.

Yong, S.Z., Zhu, M., and Frazzoli, E. (2015). Resilient state estimation against switching attacks on stochastic cyber-physical systems. In *IEEE CDC*, 5162–5169.

APPENDIX

$$M^\iota = \underset{i,j}{\mathrm{diag}}\{M_i\}, \Gamma^\iota_u = \underset{i,j}{\mathrm{vec}}\{\Gamma_{u,i}\}, \Gamma^\iota_d = \underset{i,j}{\mathrm{diag}}\{\Gamma_{d,i}\},$$
$$\Gamma^\iota_w = \underset{i,j}{\mathrm{diag}}\{\Gamma_{w,i}\}, F^\iota_T = \underset{i,j}{\mathrm{vec}}\{F_i\}, E^\iota = \underset{i,j}{\mathrm{diag}}\{E_i\},$$
$$F^\iota_u = \underset{i,j}{\mathrm{vec}}\{F_{u,i}\}, F^\iota_d = \underset{i,j}{\mathrm{diag}}\{F_{d,i}\}, F^\iota_v = \underset{i,j}{\mathrm{diag}}\{F_{v,i}\}, G^\iota = \underset{i,j}{\mathrm{vec}}\{G_i\},$$
$$R^\iota_1 = \begin{bmatrix} M_i & 0 & \Gamma_{d,i} & 0 & \Gamma_{w,i} & 0 & 0 & 0 \\ 0 & M_j & 0 & \Gamma_{d,j} & 0 & \Gamma_{w,j} & 0 & 0 \end{bmatrix},$$
$$R^\iota_2 = \begin{bmatrix} E_i & -E_j & F_{d,i} & -F_{d,j} & 0 & 0 & F_{v,i} & -F_{v,i} \\ -E_i & E_j & -F_{d,i} & F_{d,j} & 0 & 0 & -F_{v,i} & F_{v,i} \end{bmatrix},$$
$$r^\iota_1 = \begin{bmatrix} -F_i \\ -F_j \end{bmatrix},\ S^\iota_1 = \begin{bmatrix} -\Gamma_{u,i} \\ -\Gamma_{u,j} \end{bmatrix}, r^\iota_2 = \begin{bmatrix} G_j - G_i \\ G_i - G_j \end{bmatrix}, S^\iota_2 = \begin{bmatrix} F_{u,j} - F_{u,i} \\ F_{u,i} - F_{u,j} \end{bmatrix}.$$
$$M_i = \begin{bmatrix} \underline{A}_i & -\mathbb{I} & 0 & 0 & \cdots & 0 \\ -\overline{A}_i & \mathbb{I} & 0 & 0 & \cdots & 0 \\ 0 & \underline{A}_i & -\mathbb{I} & 0 & \cdots & 0 \\ 0 & -\overline{A}_i & \mathbb{I} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \underline{A}_i & -\mathbb{I} \\ 0 & 0 & \cdots & \cdots & -\overline{A}_i & \mathbb{I} \end{bmatrix}, F_i = \begin{bmatrix} \underline{f}_i \\ -\overline{f}_i \\ \underline{f}_i \\ -\overline{f}_i \\ \vdots \\ \underline{f}_i \\ -\overline{f}_i \end{bmatrix},$$
$$E_i = \underset{T}{\mathrm{diag}}\{C_i\}, G_i = \underset{T}{\mathrm{vec}}\{g_i\}.$$
$$\text{For } \star = \{d, w, u\} : \Gamma_{\star,i} = \underset{T}{\mathrm{diag}}\{\begin{bmatrix} \underline{B}_{\star,i} \\ -\overline{B}_{\star,i} \end{bmatrix}\}, F_{\star,i} = \underset{T}{\mathrm{diag}}\{D_{\star,i}\}.$$