

Active Model Discrimination for Piecewise Affine Inclusion Systems

Ruochen Niu * Elikplim Gah ** Sze Zheng Yong **

* School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, USA (e-mail: niu6@asu.edu).

** Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA, USA (e-mail: {[@northeastern.edu](mailto:gah.e,s.yong)})

Abstract: This paper proposes a novel set-membership active model discrimination (AMD) algorithm for actively separating/discriminating among a set of piecewise affine inclusion systems with bounded noise. Specifically, to overcome the difficulties of dealing the integer variables in the lower/inner level of the associated bilevel optimization problem that stem from the mapping of the piecewise inclusions and subregions, we propose an alternative reformulation that moves the integer variables into the higher/outer level. This reformulation allows us to leverage Karush-Kuhn-Tucker (KKT) conditions to obtain an equivalent single level mixed-integer linear programming (MILP) problem. Moreover, in contrast to standard AMD algorithms that strictly enforces model separation, we propose a slight modification/extension that separates as many models as possible when a strict separation of all models is not possible.

Copyright © 2023 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: Active fault diagnosis; FDI for hybrid systems; Estimation and fault detection; Model validation; Robust estimation

1. INTRODUCTION

Safety is or has to be a primary design concern when integrating dynamical systems in everyday life. However, most of these systems exhibit nonlinear system dynamics, which often makes it difficult to accurately assess system performance/safety or determine their operating modes. Nonetheless, in most cases, the true system dynamics can be over-approximated by a piecewise affine inclusion model. Hence, the ability to detect and/or enforce separation between the state/output trajectory sets of different such models would provide an effective tool for model detection and identification, including for fault detection/isolation and model validation.

Literature Review. Model discrimination methods can be categorized into two classes based on the description of uncertainties—probabilistic methods, e.g., Patton et al. (2000); Kerestecioglu and Çetin (2004); Blackmore et al. (2008) and set-based methods, e.g., Nikoukhah and Campbell (2006); Scott et al. (2014). Set-based methods, which this paper focuses on, can be further categorized into passive and active methods. Passive methods generally rely on input and output measurements to falsify or validate models by identifying differences between the expected and actual outputs (Venkatasubramanian et al. (2003); Lou and Si (2009); Harirchi et al. (2016)). However, the presence of uncertainties and disturbances over extended time horizons can have detrimental effects on the accuracy and thus, the reliability of these passive methods.

By contrast, active methods design an input to assist the model discrimination, often in the context of isolating and identifying fault models. Specifically, active model discrimination (AMD) algorithms, e.g., Nikoukhah and Campbell (2006); Cheong and Manchester (2015); Tabatabaeipour (2015), seek to obtain the smallest input/excitation that has a minimal effect on the desired behavior of the system while strictly enforcing model separation. Further, approaches with assumption of zonotopic uncertainties was proposed in Scott et al. (2014); Raimondo et al. (2013). The authors in Raimondo et al. (2016); Marseglia and Raimondo (2017) used set-valued observers for achieving closed-loop AMD, while Niu et al. (2019) introduced a partition-based AMD approach to enable the use of (side) information that is revealed at run time.

The extension of the AMD framework to nonlinear systems is nontrivial. An AMD method for nonlinear models is proposed in (Singh, Ding, Ozay, and Yong, 2018) by constructing single region/non-switched affine inclusion models (i.e., uncertain affine models) to approximate the nonlinear models, which can often lead to poor model over-approximations and thus, poor AMD performance. On the other hand, the authors in Niu et al. (2022) considered the use of multi-region/switched affine abstraction models (i.e., piecewise affine inclusion models) to over-approximate the nonlinear dynamics, as well as considered metric temporal logic switching specifications. However, the proposed AMD algorithm in Niu et al. (2022) is based on (multi-)parametric optimization that can be very computationally expensive or cumbersome and may only be applicable to certain (small) problems, even with the

* This work is supported in part by NSF grant CNS-2313814, ONR grant N00014-23-1-2093 and NASA STTR grant 80NSSC21C0371.

proposed complexity reduction strategies of shortening the time horizon and partitioning the input domain.

Contribution. In this paper, we consider a novel set-membership framework for active model discrimination (AMD) to tackle the above-mentioned difficulties for designing a separating input for piecewise affine inclusion models with bounded uncertainties. In particular, our approach is optimization-based and relies on an implicit set representation, where model separation (more specifically, separation of output sets) is enforced for worst-case realizations of set-valued uncertainties. The resulting bilevel optimization problem is shown to involve integer variables in the lower/inner level that cannot be solved by most existing optimization solvers. Thus, we propose a (potentially suboptimal) reformulation that moves the integer variables from the lower/inner level to the higher/outer level of the associated bilevel optimization problem. By doing so, we can leverage Karush-Kuhn-Tucker (KKT) conditions to recast the problem as an equivalent single-level mixed-integer linear programming (MILP) problem that is more tractable and solvable using off-the-shelf solvers, e.g., Gurobi (2015).

Further, since piecewise affine inclusion models can be used to over-approximate nonlinear dynamics up to a desired precision, the proposed algorithms are also applicable for discriminating among nonlinear models. In addition, we propose a slight modification/extension to standard AMD algorithms in order to separate as many models as possible even when not all the models are strictly distinguishable from each other. The proposed approach in this paper will benefit various applications such as active fault diagnosis, intent identification and terrain estimation.

2. PRELIMINARIES

2.1 Notations and Definitions

Let $x \in \mathbb{R}^n$ denote a vector and $M \in \mathbb{R}^{n \times m}$ a matrix, with transpose $M^\top$ and $M \geq 0$ denotes element-wise non-negativity. The infinity vector norm of x is denoted by $\|x\| \triangleq \max_i x_i$, while $\mathbf{0}$, $\mathbf{1}$ and $\mathbb{I}$ represent the matrix 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 $M_i, i = 1, \dots, n$ and matrix M as follows:

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

where $\otimes$ is the Kronecker product. The set of positive integers up to n is denoted by $\mathbb{Z}_n^+$, and the set of integers from m to n is denoted by $\mathbb{Z}_m^n$. Further, we adopt the definition of Special Ordered Set of degree 1 (SOS-1) constraints, e.g., in Gurobi (2015), which constrains that only one element in a set of (continuous, integer or mixed-integer) variables can be non-zero, i.e., given a set $S = \{v_1, \dots, v_n\}$, if $v_i \neq 0$ for any i , then $v_j = 0, \forall j \neq i$. Such constraints are readily enforced in many off-the-shelf solvers, e.g., Gurobi (2015), and they often help to speed up mixed-integer optimization.

2.2 Modeling Framework

Consider N discrete-time piecewise affine inclusion system models $\{\mathcal{H}_\ell\}_{\ell=1}^N$, each with states $\mathbf{x}_\ell \in \mathbb{R}^n$, measurements/outputs $z_\ell \in \mathbb{R}^p$, inputs $u \in \mathbb{R}^m$, process noise $w_\ell \in \mathbb{R}^{m_w}$, and measurement noise $v_\ell \in \mathbb{R}^{m_v}$. The system dynamics for each piece/subregion $i \in \mathbb{Z}_{q_\ell}^+$ of model $\mathcal{H}_\ell$ (with its domain partitioned into q_ℓ subregions) is given by:

$$\begin{aligned} \begin{pmatrix} \underline{A}_{\ell,i} \mathbf{x}_\ell(k) + \underline{B}_{\ell,i}^u u(k) \\ + \underline{B}_{\ell,i}^w w_\ell(k) + \underline{f}_{\ell,i} \end{pmatrix} &\leq \mathbf{x}_\ell(k+1) \leq \begin{pmatrix} \overline{A}_{\ell,i} \mathbf{x}_\ell(k) + \overline{B}_{\ell,i}^u u(k) \\ + \overline{B}_{\ell,i}^w w_\ell(k) + \overline{f}_{\ell,i} \end{pmatrix}, \quad (1) \\ z_\ell(k) &= C_\ell \mathbf{x}_\ell(k) + D_\ell^u u(k) + D_\ell^v v_\ell(k) + g_\ell. \quad (2) \end{aligned}$$

with known constant matrices and vectors of appropriate dimensions, $\Phi_{\ell,i}$ and $\overline{\Phi}_{\ell,i}$ for all $\Phi \in \{A, B^u, B^w, f\}$, where $\mathbf{x}_\ell(k+1)$ denotes the state at the next time instant, and their corresponding polytopic subregions $I_{\ell,i}$ is given by the following linear constraints:

$$S_{\ell,i}^x \mathbf{x}_\ell(k) + S_{\ell,i}^u u(k) + S_{\ell,i}^w w_\ell(k) + \beta_{\ell,i} \leq 0, \quad (3)$$

with $S_{\ell,i}^x, S_{\ell,i}^u, S_{\ell,i}^w, \beta_{\ell,i}$ of appropriate dimensions. Note the above piecewise affine inclusion models could also represent an over-approximation/abstraction of nonlinear models that can be obtained via affine abstraction methods in (Singh, Shen, and Yong, 2018); Jin et al. (2020).

The states $\mathbf{x}_\ell$ can be divided into controlled states $x_\ell \in \mathbb{R}^{n_x}$ and uncontrolled states $y_\ell \in \mathbb{R}^{n_y}$ with $n_y = n - n_x$ accordingly. As a consequence, we have

$$\mathbf{x}_\ell(k) \triangleq [x_\ell(k)^\top \ y_\ell(k)^\top]^\top. \quad (4)$$

The initial condition for model $\mathcal{H}_\ell$, denoted by $\mathbf{x}_\ell^0 = \mathbf{x}_\ell(0)$, is constrained to a polyhedral set with c_0 inequalities:

$$\mathbf{x}_\ell^0 \in \mathcal{X}_0 := \{\mathbf{x} \in \mathbb{R}^n : P_0 \mathbf{x} \leq p_0\}, \forall \ell \in \mathbb{Z}_N^+. \quad (5)$$

Moreover, the states x_ℓ and y_ℓ satisfy polyhedral state constraints with c_x and c_y inequalities:

$$x_\ell(k) \in \mathcal{X}_{x,\ell} := \{x \in \mathbb{R}^{n_x} : P_{x,\ell} x \leq p_{x,\ell}\}, \quad (6)$$

$$y_\ell(k) \in \mathcal{X}_{y,\ell} := \{y \in \mathbb{R}^{n_y} : P_{y,\ell} y \leq p_{y,\ell}\}. \quad (7)$$

On the other hand, the controlled inputs u must also satisfy the following polyhedral constraints with c_u inequalities:

$$u(k) \in \mathcal{U} := \{u \in \mathbb{R}^{m_u} : Q_u u \leq q_u\}. \quad (8)$$

The process and measurement noises, w_ℓ and v_ℓ , are also polyhedrally constrained with c_w and c_y inequalities:

$$w_\ell(k) \in \mathcal{W}_\ell := \{w \in \mathbb{R}^{m_w} : Q_{w,\ell} w \leq q_{w,\ell}\}, \quad (9)$$

$$v_\ell(k) \in \mathcal{V}_\ell := \{v \in \mathbb{R}^{m_v} : Q_{v,\ell} v \leq q_{v,\ell}\}. \quad (10)$$

The readers are referred to [Remark 1, Ding et al. (2018)] for a description of the well-posedness of the formulation.

2.3 Matrices Concatenation

Given time horizon T , we first introduce some useful time-concatenated notations for system variables/signals. The time-concatenated states and outputs are as follows:

$$\mathbf{x}_{\ell,T} = \text{vec}_{k=0}^T \{\mathbf{x}_\ell(k)\}, \quad z_{\ell,T} = \text{vec}_{k=0}^T \{z_\ell(k)\},$$

while the time-concatenated inputs and noises are

$$\begin{aligned} u_T &= \text{vec}_{k=0}^{T-1} \{u(k)\}, \quad w_{\ell,T} = \text{vec}_{k=0}^{T-1} \{w_\ell(k)\}, \\ v_{\ell,T} &= \text{vec}_{k=0}^{T-1} \{v_\ell(k)\}. \end{aligned}$$

Using the above time-concatenated inputs, noise, states and outputs, the corresponding time-concatenated state and subregion inequalities in (1) and (3) as well as the output equation in (2) can be written as:

$$M_{\ell,i} \mathbf{x}_{\ell,T} + \Gamma_{\ell,i}^u u_T + \Gamma_{\ell,i}^w w_{\ell,T} + F_{\ell,i,T} + \alpha_{\ell,i} \leq 0, \quad (11)$$

$$\Omega_{\ell,i}^x \mathbf{x}_{\ell,T} + \Omega_{\ell,i}^u u_T + \Omega_{\ell,i}^w w_{\ell,T} + \beta_{\ell,i,T} + \zeta_{\ell,i} \leq 0, \quad (12)$$

$$z_{\ell,T} = E_{\ell} \mathbf{x}_{\ell,T} + F_{\ell}^u u_T + F_{\ell}^v v_{\ell,T} + G_{\ell}, \quad (13)$$

for all $\ell \in \mathbb{Z}_N^+$, $i \in \mathbb{Z}_{q_\ell}^+$, where $\alpha_{\ell,i} = \text{vec}_{k=0}^{T-1} \{\text{vec}_{2n} \{s_{\ell,i}(k)\}\}$ and $\zeta_{\ell,i} = \text{vec}_{k=0}^{T-1} \{\text{vec}_{2p} \{s_{\ell,i}(k)\}\}$. $s_{\ell,i}(k)$ is the unconstrained continuous slack variable for each piece/subregion i . Specifically, when $s_{\ell,i}(k) = 0$, the subregion i is valid at k time step, while when $s_{\ell,i}(k)$ is free/unconstrained, the subregion i and corresponding inequalities in (11) and (12) hold trivially. Thus, for each $k \in \mathbb{Z}_{T-1}^0$, a piecewise affine inclusion model with the state in subregion i has to satisfy:

$$s_{\ell,i}(k) = 0, s_{\ell,j}(k) \neq 0, \forall i, j \in \mathbb{Z}_{q_\ell}^+, i \neq j. \quad (14)$$

We explicitly encode the above to enforce that only one subregion/piece is valid at each time step, by introducing a binary variable $a_{\ell,i}(k)$ that satisfies the following:

$$\forall \ell \in \mathbb{Z}_N^+, \forall k \in \mathbb{Z}_T^0, \forall i \in \mathbb{Z}_{q_\ell}^+, : a_{\ell,i}(k) \in \{0, 1\}, \quad (15a)$$

$$\sum_{i=1}^{q_\ell} a_{\ell,i}(k) = 1, \text{SOS-1:} \{a_{\ell,i}(k), s_{\ell,i}(k)\}. \quad (15b)$$

All the matrices and vectors in (11)–(13) are defined in the Appendix with suitable dimensions. Further, the uncertain variables for each model ℓ are concatenated as $\bar{x}^\ell = \text{vec}\{\mathbf{x}_{\ell,T}, w_{\ell,T}\}$.

Then, given N discrete-time piecewise affine inclusion models, there are $\mathcal{I} = \binom{N}{2}$ model pairs and let the mode $\iota \in \{1, \dots, \mathcal{I}\}$ denote the pair of models $(\dagger, \ddagger)$. Then, we define the pair-concatenated state, output and noises:

$$\mathbf{x}_T^\iota = \text{vec}_{\ell=\{\dagger, \ddagger\}} \{\mathbf{x}_{\ell,T}\}, z_T^\iota = \text{vec}_{\ell=\{\dagger, \ddagger\}} \{z_{\ell,T}\},$$

$$w_T^\iota = \text{vec}_{\ell=\{\dagger, \ddagger\}} \{w_{\ell,T}\}, v_T^\iota = \text{vec}_{\ell=\{\dagger, \ddagger\}} \{v_{\ell,T}\}.$$

The states, corresponding subregions and outputs over the entire time horizon in (11), (12) and (13) for each pair ι can be written as simple functions of the state $\mathbf{x}_T^\iota$, input vectors u_T , and noise w_T^ι, v_T^ι :

$$M^\iota \mathbf{x}_T^\iota + \Gamma_u^\iota u_T + \Gamma_w^\iota w_T^\iota + F_T^\iota \leq 0, \quad (16)$$

$$z_T^\iota = E^\iota \mathbf{x}_T^\iota + F_u^\iota u_T + F_v^\iota v_T^\iota + G^\iota, \quad (17)$$

where the matrices and vectors $M^\iota, \Gamma_u^\iota, \Gamma_w^\iota, F_T^\iota, E^\iota, F_u^\iota, F_v^\iota$ and G^ι are defined in the Appendix. Moreover, the uncertain variables for each pair ι are concatenated as $\bar{x}^\iota = \text{vec}\{\mathbf{x}_T^\iota, w_T^\iota, v_T^\iota\}$.

We then concatenate the polyhedral state constraints in (6) and (7) in terms of time and pair. For $\hat{x} \in \{x, y\}$, let

$$\bar{P}_{\hat{x},\ell} = \text{diag}_T \{P_{\hat{x},\ell}\}, \bar{p}_{\hat{x},\ell} = \text{vec}_T \{p_{\hat{x},\ell}\},$$

$$\bar{P}_{\hat{x}} = [\text{diag}_{\ell=\{\dagger, \ddagger\}} \{\bar{P}_{\hat{x},\ell}\} \mathbf{0}], \bar{p}_{\hat{x}} = \text{vec}_{\ell=\{\dagger, \ddagger\}} \{\bar{p}_{\hat{x},\ell}\}.$$

Then, we can rewrite the polyhedral constraints as:

$$\bar{P}_{\hat{x}}^\iota \bar{x}^\iota \leq \bar{p}_{\hat{x}}^\iota.$$

Further, the initial state constraint in (5) can be written as $\bar{P}_0^\iota \bar{x}^\iota \leq \bar{p}_0^\iota$ with

$$\bar{P}_0^\iota = \text{diag}_2 \{[P_0 \mathbf{0}]\}, \bar{p}_0^\iota = \text{vec}_2 \{p_0\}.$$

Similarly, let

$$\bar{Q}_u = \text{diag}_T \{Q_u\}, \bar{Q}_{\dagger}^\iota = \text{diag}_{\ell=\{\dagger, \ddagger\}} \text{diag}_T \{Q_{\dagger,\ell}\},$$

$$\bar{q}_u = \text{vec}_T \{q_u\}, \bar{q}_{\dagger}^\iota = \text{vec}_{\ell=\{\dagger, \ddagger\}} \text{vec}_T \{q_{\dagger,\ell}\}, \dagger \in \{w, v\}.$$

Then, the input constraints in (8) and (9) for all k are: $\bar{Q}_u u_T \leq \bar{q}_u$ and $\bar{Q}_{\dagger}^\iota \dagger_T^\iota \leq \bar{q}_{\dagger}^\iota$. Hence, in terms of $\bar{x}^\iota$, we have a polyhedral constraint of the form $H_{\bar{x}}^\iota \bar{x}^\iota \leq h_{\bar{x}}^\iota$, with

$$H_{\bar{x}}^\iota = \begin{bmatrix} \bar{P}_0^\iota & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{P}_y^\iota & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \text{diag}\{\bar{Q}_w^\iota, \bar{Q}_v^\iota\} \end{bmatrix},$$

$$h_{\bar{x}}^\iota = \text{vec}\{\bar{p}_0^\iota, \bar{p}_y^\iota, \bar{q}_w^\iota, \bar{q}_v^\iota\}.$$

Note that the above definitions of matrices and vectors for the *piecewise affine inclusion model* in (1) are different from the ones in Ding et al. (2018); (Singh, Ding, Ozay, and Yong, 2018) that were defined for *affine (inclusion) models* (i.e., a special case of our considered model in (1) with only one piece/subregion).

3. PROBLEM FORMULATION

The goal of this paper is to design an AMD algorithm for piecewise affine inclusions, i.e., we look to find an admissible input for the models such that this input causes the reachable output set of any pair of models to have no intersection for at least one time instant over the entire time horizon T , considering all possible realizations of set-valued uncertainties.

Problem 0 (Active Model Discrimination for $\{\mathcal{H}_\ell\}_{\ell=1}^N$). *Given N piecewise affine inclusion models $\{\mathcal{H}_\ell\}_{\ell=1}^N$, and state, input and noise constraints, i.e., (5)–(10), find an optimal input sequence u_T^* to minimize a given cost function $J(u_T)$ such that for all possible initial states $\mathbf{x}_\ell^0$, process noise $w_{\ell,T}$ and measurement noise $v_{\ell,T}$, only one model is valid, i.e., the output trajectories of any pair of models have to differ by a threshold ϵ in at least one time step. The optimization problem can be formally stated as:*

$$u_T^* = \arg \min_{u_T} J(u_T)$$

$$\text{s.t.} \quad \forall k \in \mathbb{Z}_{T-1}^0 : (8) \text{ holds}, \quad (18a)$$

$$\forall \mathbf{x}_\ell^0, w_{\ell,T}, \alpha_{\ell,i}, \zeta_{\ell,i}, a_{\ell,i}(k) : (6) \text{ holds}, \quad (18b)$$

$$\forall \mathbf{x}_\ell^0, w_T^\iota, v_T^\iota, \alpha_\ell, \zeta_\ell, a_{\ell,i}(k) : \{\exists k' \in \mathbb{Z}_{T-1}^0 : (5), (7), (9), (10), (15) \text{--}(17) \text{ hold.}\} \implies \|z_{\dagger}^\iota(k') - z_{\ddagger}^\iota(k')\| \geq \epsilon, \quad (18c)$$

where $J(\cdot)$ is a convex objective function of the inputs u_T .

Note that the above problem is similar to the one considered in Niu et al. (2022). The key difficulty of solving this problem is the presence of the constraint (15) and the corresponding integer variables $a_{\ell,i}(k)$ on the left-hand side of (18b) and (18c) that leads to binary variables/constraints in the lower/inner level of any resulting bilevel optimization problem formulations. In Niu et al. (2022), this problem is directly solved but requires the use of (multi-)parametric optimization subroutines for mixed-integer linear problems, which can be computationally expensive and is often restricted to very small problems. Thus, Niu et al. (2022) also proposed strategies for reducing the computational time. Even so, the problem is often still computational expensive and limited to smaller problems. In this paper, we propose to solve an alternative reformulation, given in the problem below, that does not require (multi-)parametric programming; thus, it is faster and more tractable, but with a slight loss of optimality.

Moreover, we also propose a slight modification/extension of standard AMD problems (that strictly enforces separa-

tion of all model pairs) to separate as many model pairs as possible, which can be useful in certain settings where not all model pairs are distinguishable. To achieve this, the problem below introduces a binary variable b^ι for each model pair indexed by $\iota \in \mathbb{Z}_T^+$ to indicate the successful discrimination for each pair and penalize the objective function to separate as many pairs of models as possible, i.e., $b^\iota = 1$ implies the successful discrimination of the ι -th model pair $(\dagger, \ddagger)$ with at least ϵ , which is equivalent to

$$\forall \iota \in \mathbb{Z}_T^+, \exists k \in \mathbb{Z}_T^0 : \|z_\dagger(k) - z_\ddagger(k)\| \geq \epsilon - \bar{s}^\iota, \quad (19a)$$

$$\text{SOS-1} : \{b^\iota, \bar{s}^\iota\}, b^\iota \in \{0, 1\}, \quad (19b)$$

where $\bar{s}^\iota \in \mathbb{R}$ is an unconstrained/free continuous slack variable. $b^\iota = 1$ enforces that $\bar{s}^\iota = 0$, and consequently, (19a) holds with $\bar{s}^\iota = 0$ (i.e., the output sets differ by a ‘distance’ of at least ϵ). Otherwise, if $b^\iota = 0$, $\bar{s}^\iota$ is unconstrained and (19a) trivially holds.

The reformulated AMD problem is then as shown in Problem 1. Note that the reformulation is such that its optimal solution is a feasible (but potentially suboptimal) solution to Problem 0.

Problem 1 (Active Model Discrimination for $\{\mathcal{H}_\ell\}_{\ell=1}^N$). *Given N well-posed piecewise affine inclusion models $\{\mathcal{H}_\ell\}_{\ell=1}^N$, and state, input and noise constraints, i.e., (5)–(10), such that for all possible initial states $\mathbf{x}_\ell^0$, process noise $w_{\ell,T}$ and measurement noise $v_{\ell,T}$, the models are separated as much as possible, i.e., the output trajectories of as many model pairs as possible have to differ by a threshold ϵ in at least one time instance. The optimization problem can be formally written as:*

$$u_T^* = \arg \min_{u_T, b^\iota} J(u_T) - \lambda \sum_{\iota=1}^T b^\iota \quad (20a)$$

$$\text{s.t.} \quad \forall k \in \mathbb{Z}_{T-1}^0 : (8) \text{ holds}, \quad (20a)$$

$$\forall k \in \mathbb{Z}_T^0, \forall \ell \in \mathbb{Z}_N^+, \forall i \in \mathbb{Z}_{q_\ell}^+ : (15) \text{ holds}, \quad (20b)$$

$$\forall \mathbf{x}_\ell^0, w_\ell^t, v_\ell^t, \alpha_\ell, \zeta_\ell : \{ (5), (7), (9), (11) \text{ hold} \} : (6), (12) \text{ hold}, \quad (20c)$$

$$\forall \mathbf{x}_\ell^0, w_\ell^t, v_\ell^t, \alpha_\ell, \zeta_\ell : \{ (5), (7), (9), (10), (16), (17) \text{ hold} \} : (19) \text{ holds}, \quad (20d)$$

where λ is a constant tuning parameter¹ for penalizing the binary variables b^ι .

In contrast to Problem 0, the key change in our alternative reformulation in Problem 1 is to move the ‘problematic’ constraint (12), (15) and the corresponding integer variables $a_{\ell,i}(k)$ to the right hand side of (20b) and (20c). In doing so, we allow the designer of the separating input (in the upper/outer level) to take on the ‘responsibility’ for selecting the subregions (via (15) that enforces (14)) and for enforcing that the selected subregions (via (12)) must be satisfied. Thus, (20b) and (20c) both guarantee that the system dynamics remains in the selected subregion and the controlled state constraints are satisfied, in spite of all possible realizations of uncertainties over the entire time horizon for each model. However, by requiring that the designer of the separating input to satisfy this additional ‘responsibility,’ the solution to Problem 1 is only sufficient (and potentially suboptimal) for solving Problem 0.

Remark 1. Since the piecewise affine inclusion models described by (1)–(3) can also represent an over-approximation/abstraction of nonlinear models (algorithms

for automating the abstraction process can be found in (Singh, Shen, and Yong, 2018); Jin et al. (2020)), the AMD algorithm we propose for solving Problem 1 can also be applied to solve the active discrimination problem for nonlinear models.

4. MAIN RESULT

In this section, we first formulate Problem 1 as a bilevel optimization problem, which then is recast as an equivalent single-level MILP by applying KKT conditions to the lower/inner problem (where there are no longer any integer variables). The bilevel formulation is as follows:

Lemma 1 (Bilevel Optimization Formulation). *Given a separability index ϵ , the active model discrimination problem in Problem 1 is equivalent to a bilevel optimization problem with the following outer problem:*

$$u_T^* = \arg \min_{u_T, b^\iota} J(u_T) - \lambda \sum_{\iota=1}^T b^\iota \quad (\text{P}_{Outer})$$

$$\text{s.t.} \quad \forall k \in \mathbb{Z}_{T-1}^0 : (8) \text{ holds}, \quad (21a)$$

$$\forall k \in \mathbb{Z}_T^0, \forall \ell \in \mathbb{Z}_N^+, \forall i \in \mathbb{Z}_{q_\ell}^+ : (15) \text{ holds}, \quad (21b)$$

$$\forall \mathbf{x}_\ell^0, w_\ell^t, v_\ell^t, \alpha_\ell, \zeta_\ell : \{ (5), (7), (9), (11) \text{ hold} \} : (6), (12) \text{ hold}, \quad (21c)$$

$$\forall \iota \in \mathbb{Z}_T^+ : \delta^{\iota*}(u_T) \geq \epsilon - \bar{s}^\iota, \quad (21d)$$

where $\delta^{\iota*}$ is the solution of the inner problem:

$$\delta^{\iota*}(u_T) = \min_{\delta^\iota, \mathbf{x}_T^\iota, w_T^\iota, v_T^\iota} \delta^\iota \quad (\text{P}_{Inner})$$

$$\text{s.t.} \quad \forall \iota \in \mathbb{Z}_T^+, k \in \mathbb{Z}_T^0 : (16), (17) \text{ hold}, \quad (22a)$$

$$\forall \iota \in \mathbb{Z}_T^+, k \in \mathbb{Z}_T^0 : \|z_\dagger(k) - z_\ddagger(k)\| \leq \delta^\iota, \quad (22b)$$

$$\forall \mathbf{x}_\ell^0, w_\ell^t, v_\ell^t : (5), (7), (9), (10) \text{ hold}, \quad (22c)$$

with the ι -th model pair being $(\dagger, \ddagger)$ where $\dagger, \ddagger \in \mathbb{Z}_N^+, \dagger \neq \ddagger$.

Proof. The proof follows a similar procedure as in Ding et al. (2018). $\square$

This bilevel problem can then be recast into a single-level MILP that can be readily solved using off-the-shelf solvers, given as follows:

Theorem 1 (Active Model Discrimination for $\{\mathcal{H}_\ell\}_{\ell=1}^N$). *Given a separability index ϵ , the bilevel optimization problem in Lemma 1 is equivalent to the following mixed-integer optimization problem:*

$$u_T^* = \arg \min_{u_T, \bar{x}^\ell, \nu_1^\ell, \nu_2^\ell, a_{\ell,i}(k), s_{\ell,i}(k), \bar{x}^\iota, \mu_1^\iota, \mu_2^\iota, \mu_3^\iota, b^\iota, \bar{s}^\iota, \delta^\iota} J(u_T) - \lambda \sum_{\iota=1}^T b^\iota$$

$$\text{s.t.} \quad \bar{Q}_u u_T \leq \bar{q}_u, \quad (23)$$

$$\forall \ell \in \mathbb{Z}_N^+, \forall a \in \mathbb{Z}_{c_x}^+ : (32a) - (32e) \text{ hold}, \quad (24)$$

$$\forall \iota \in \mathbb{Z}_T^+ : \begin{cases} \delta^\iota + \bar{s}^\iota \geq \epsilon, \\ b^\iota \in \{0, 1\}, \text{ SOS-1} : \{b^\iota, \bar{s}^\iota\}, \\ 0 = 1 - \mu_3^\iota \mathbf{T} \mathbf{1}, \\ 0 = \sum_{i=1}^{i=\kappa} \mu_{1,i}^\iota H_{\bar{x}}^\iota(i, m) + \sum_{j=1}^{j=\xi} \mu_{2,j}^\iota R_1^\iota(j, m) \\ \quad + \sum_{k=1}^{k=\rho} \mu_{3,k}^\iota R_2^\iota(k, m), \forall m = 1, \dots, \eta, \\ \tilde{H}_{\bar{x},i}^\iota \bar{x}^\iota - h_{\bar{x},i}^\iota \leq 0, \mu_{1,i}^\iota \geq 0, \forall i = 1, \dots, \kappa, \\ \tilde{R}_{1,j}^\iota \bar{x}^\iota - r_{1,j}^\iota - S_{1,j}^\iota u_T \leq 0, \\ \mu_{2,j}^\iota \geq 0, \forall j = 1, \dots, \xi, \\ \tilde{R}_{2,k}^\iota \bar{x}^\iota - \delta^\iota - r_{2,k}^\iota - S_{2,k}^\iota u_T \leq 0, \mu_{3,k}^\iota \geq 0, \\ \forall k = 1, \dots, \rho, \end{cases} \quad (25)$$

¹ For instance, a large λ (e.g., $\lambda = 100$) will incentivize separation of as many model pairs as possible over minimizing control effort.

$$\forall \ell \in \mathbb{Z}_N^+, \forall m \in \mathbb{Z}_\alpha^+ : \text{SOS-1} : \{\nu_{1,a,m}^\ell, \tilde{P}_{x,m}^\ell \tilde{x}_a^\ell - \tilde{P}_{u,m}^\ell u_T - \tilde{p}_m^\ell\}, \quad (26)$$

$$\forall \ell \in \mathbb{Z}_N^+, \forall n \in \mathbb{Z}_\gamma^+ : \text{SOS-1} : \{\nu_{2,a,n}^\ell, \tilde{H}_n^\ell \tilde{x}_a^\ell - \tilde{h}_n^\ell\}, \quad (27)$$

$$\forall \ell \in \mathbb{Z}_T^+, \forall i \in \mathbb{Z}_\kappa^+ : \text{SOS-1} : \{\mu_{1,i}^\ell, \tilde{H}_{x,i}^\ell \tilde{x}^\ell - \tilde{h}_{x,i}^\ell\}, \quad (28)$$

$$\forall \ell \in \mathbb{Z}_T^+, \forall j \in \mathbb{Z}_\xi^+ : \text{SOS-1} : \{\mu_{2,j}^\ell, \tilde{R}_{1,j}^\ell \tilde{x}^\ell - r_{1,j}^\ell - \tilde{S}_{1,j}^\ell u_T\}, \quad (29)$$

$$\forall \ell \in \mathbb{Z}_T^+, \forall k \in \mathbb{Z}_\rho^+ : \text{SOS-1} : \{\mu_{3,k}^\ell, \tilde{R}_{2,k}^\ell \tilde{x}^\ell - \delta^\ell - r_{2,k}^\ell - \tilde{S}_{2,k}^\ell u_T\}, \quad (30)$$

$$\begin{aligned} \forall \ell \in \mathbb{Z}_N^+, \forall i \in \mathbb{Z}_{q_\ell}^+ : \text{SOS-1} : \{a_{\ell,i}(k), s_{\ell,i}(k)\} \\ \forall k \in \mathbb{Z}_{T-1}^0 : \{a_{\ell,i}(k) \in \{0, 1\}, \sum_{i=1}^{q_\ell} a_{\ell,i}(k) = 1\}, \end{aligned} \quad (31)$$

where $\nu_{1,a,m}^\ell$, $\nu_{2,a,n}^\ell$, $\mu_{1,i}^\ell$, $\mu_{2,j}^\ell$ and $\mu_{3,k}^\ell$ are dual variables from KKT conditions, while $\tilde{P}_{*,m}^\ell$ is the m -th row of $\tilde{P}_*^\ell$, $*$ $\in \{x, u\}$, $\tilde{H}_n^\ell$ is the n -th row of $\tilde{H}^\ell$, $\tilde{H}_{x,i}^\ell$ is the i -th row of $\tilde{H}_x^\ell$, $\tilde{R}_{1,j}^\ell$ and $\tilde{S}_{1,j}^\ell$ are the j -th row of R_1^ℓ and S_1^ℓ , respectively, $\tilde{R}_{2,k}^\ell$ and $\tilde{S}_{2,k}^\ell$ are the k -th row of R_2^ℓ and S_2^ℓ , respectively, $\eta = 2IT(n + m_w + m_v)$ is the number of columns of $\tilde{H}_x^\ell$, $\kappa = 2IT(c_0 + c_w + c_v)$ is the number of rows of $\tilde{H}_x^\ell$, $\xi = 2IT(c_x + c_y)$ is the number of rows of R_1^ℓ and $\rho = 2ITp$ is the number of rows of R_2^ℓ .

Proof. We formulate the responsibility of the controlled input in (20c) as a semi-infinite constraint that can be written as:

$$\tilde{P}_x^\ell \tilde{x}^\ell \leq \tilde{P}_u^\ell u_T + \tilde{p}^\ell, \forall \tilde{x}^\ell \in \left\{ \tilde{x}^\ell : \tilde{P}_x^\ell \tilde{x}^\ell \leq \tilde{P}_u^\ell u_T + \tilde{p}^\ell, \tilde{H}^\ell \tilde{x}^\ell \leq \tilde{h}^\ell \right\}.$$

From (Shen et al., 2022, Lemma 1, Section 4), we can derive the equivalent mixed-integer constraints as:

$$0 = \sum_{m=1}^{m=\alpha} \nu_{1,a,m}^\ell \tilde{P}_x^\ell(m, \dagger) + \sum_{n=1}^{n=\gamma} \nu_{2,a,n}^\ell \tilde{H}^\ell(n, \dagger) - \tilde{P}_x^\ell(a, \dagger), \quad \forall \dagger = 1, \dots, \psi, \quad (32a)$$

$$\pi_a = \tilde{P}_{x,a}^\ell \tilde{x}_a^\ell - \tilde{P}_{u,a}^\ell u_T - \tilde{p}_a^\ell, \pi_a \leq 0, \quad (32b)$$

$$\tilde{P}_{x,m}^\ell \tilde{x}_a^\ell - \tilde{P}_{u,m}^\ell u_T - \tilde{p}_m^\ell \leq 0, \quad (32c)$$

$$\nu_{1,a,m}^\ell \geq 0, \forall m = 1, \dots, \alpha, \quad (32d)$$

$$\tilde{H}_n^\ell \tilde{x}_a^\ell - \tilde{h}_n^\ell \leq 0, \nu_{2,a,n}^\ell \geq 0, \forall n = 1, \dots, \gamma, \quad (32e)$$

$$\nu_{1,a,m}^\ell (\tilde{P}_{x,m}^\ell \tilde{x}_a^\ell - \tilde{P}_{u,m}^\ell u_T - \tilde{p}_m^\ell) = 0, \quad (32f)$$

$$\nu_{2,a,n}^\ell (\tilde{H}_n^\ell \tilde{x}_a^\ell - \tilde{h}_n^\ell) = 0, \quad (32g)$$

where $\nu_{1,a,m}^\ell$, $\nu_{2,a,n}^\ell$ are dual variables from the KKT conditions, while $\tilde{P}_{*,m}^\ell$ is the m -th row of $\tilde{P}_*^\ell$, $*$ $\in \{x, u\}$, $\tilde{H}_n^\ell$ is the n -th row of $\tilde{H}^\ell$. The complementary slackness constraints given in (32f) and (32g) are imposed using SOS-1 conditions in (27) and (28), respectively.

Furthermore, consider $a_{\ell,i^*}(k) = 1$ for some $k \in \mathbb{Z}_{T-1}^0$, then the additional constraint $\sum_{i=1}^{q_\ell} a_{\ell,i}(k) = 1$ ensures that $a_{\ell,i}(k) = 0$ for all $i \neq i^*$. Then because of the SOS-1 constraints, it follows that $s_{\ell,i}(k)$ is unconstrained for all $i \neq i^*$. Since this holds for any $k \in \mathbb{Z}_{T-1}^0$, then for each model pair, the set of feasible states for both models lies within the same subregion.

To obtain the remaining parts of the MILP formulation, we refer the reader to similar steps in (Ding et al., 2018, Section IV). $\square$

5. SIMULATION EXAMPLES

In this section, we apply the proposed approach to a lunar/planetary assisted drive system and a vehicle lane changing scenario. Both examples are implemented in MATLAB 2019b with Gurobi v9.5. Note the vehicle lane changing example uses a sparse notation to conserve space.

5.1 Lunar/Planetary Assisted Drive

First, we consider a parameter identification problem for a lunar/planetary rover navigating an environment with uncertain terrain parameters. The rover dynamics evolves according to a simple viscoelastic model with the damping coefficient being the uncertain parameter, as follows:

$$x(k+1) = x(k) + v_x(k)\delta t,$$

$$v_x(k+1) = -\frac{k}{M}x(k) + \left(1 - \frac{c}{M}\delta t\right)v_x(k) + \frac{\delta t}{M}u(k) + w(k)\delta t,$$

where x and v_x are the position and velocity of the rover in m and $\frac{m}{s}$ respectively, u is the acceleration input in $\frac{m}{s^2}$, w is the process noise signal in $\frac{m}{s^2}$, and M , k , and c are the mass, spring constant, and damping coefficient in kg , $\frac{N}{m}$, and $\frac{Ns}{m}$ respectively. δt is the sampling time in s . For this example, we have:

$$u(k) \in [-20, 20] \frac{m}{s^2}, \quad \delta t = 0.5s, \quad k = 0.002 \frac{N}{m},$$

$$w(k) \in [-0.001, 0.001] \frac{m}{s^2}, \quad M = 5kg.$$

We consider 3 uncertain models defined such that each model describes the robot dynamics given $c \in \mathcal{C}_\ell$, where $\mathcal{C}_\ell = [c_{l,\ell}, c_{u,\ell}]$. For each model, we abstracted/over-approximated the uncertain linear models, Jin et al. (2020), for each Model $\ell \in \{1, 2, 3\}$ with $c_{l,1} = 0.05$, $c_{u,1} = 0.2$, $c_{l,2} = 0.3$, $c_{u,2} = 0.5$, $c_{l,3} = 0.6$, and $c_{u,3} = 0.8$ to yield a piecewise affine inclusion model with two subregions that describes the rover trajectories:

$$\begin{aligned} \bar{A}_{\ell,2} = \bar{A}_{\ell,1} = \underline{A}_\ell, \quad \bar{A}_{\ell,1} = \bar{A}_{\ell,2} = \bar{A}_\ell, \\ B_\ell^u = B_\ell^w = \begin{bmatrix} 0 \\ \delta t \\ m \end{bmatrix}, \quad C_\ell = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_\ell^u = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad D_\ell^v = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ S_{\ell,1}^x = S_{\ell,2}^x = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}, \quad \beta_{\ell,1} = \begin{bmatrix} 180 \\ 90 \\ 180 \\ 0 \end{bmatrix}, \quad \beta_{\ell,2} = \begin{bmatrix} 180 \\ 0 \\ 180 \\ 90 \end{bmatrix}, \end{aligned}$$

where $\bar{A}_\ell \triangleq \begin{bmatrix} 1 & \delta t \\ -\frac{k}{m} & (1 - \frac{c_{u,\ell}}{m}\delta t) \end{bmatrix}$, $\underline{A}_\ell \triangleq \begin{bmatrix} 1 & \delta t \\ -\frac{k}{m} & (1 - \frac{c_{l,\ell}}{m}\delta t) \end{bmatrix}$. The system output is given by:

$$z_\ell(k) = C_\ell \begin{bmatrix} x(k) \\ v_x(k) \end{bmatrix} + D_\ell^u u(k) + D_\ell^v v(k)$$

with measurement noise $v(k) = [v_1(k) \ v_2(k)]^\top$ with $v_i(k) \in [-0.001, 0.001]$, for all $i \in \{1, 2\}$. The system is initialized with: $x(0) = 0$ and $v_x(0) \in [0.5, 0.6]$.

Solving the AMD problem with $J(\cdot) = \|\cdot\|$ and $\lambda = 100$ yields $\{u(k)\}_{k=0}^4 = \{20, 20, 20, 14.0864, 0\}$. The result from executing this input is shown in Figure 1, which shows the output reachable sets of each model. Here, we see that the output reachable sets are disjoint and thus, a system observer/operator with access to the agent's trajectory could identify the true damping coefficient range of the lunar/planetary terrain.

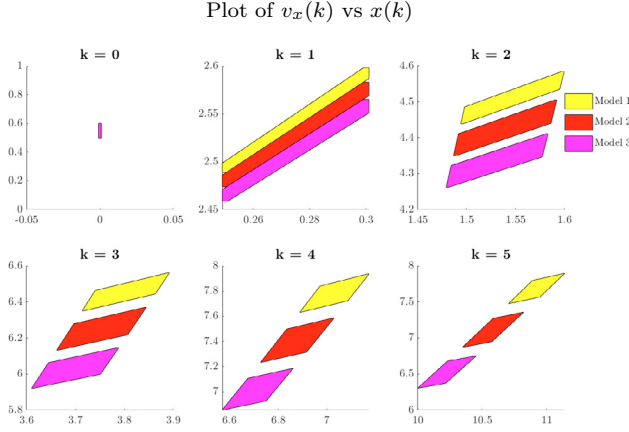


Fig. 1. Separated output sets corresponding to 3 different regions of friction coefficient, $l \in \{1, 2, 3\}$, with $\epsilon = 0.01$.

5.2 Vehicle Lane Changing Example

Similar to (Singh, Ding, Ozay, and Yong, 2018), we consider an intent estimation problem in a vehicle lane changing scenario with three intent models for the other car $\ell \in \{I, C, M\}$ (see (Singh, Ding, Ozay, and Yong, 2018) for details): the *Inattentive* driver is unaware of the ego car and maintains his speed, the *Cautious* driver tends to yield the lane to ego car or the *Malicious* driver does not want to yield the lane to the ego car. However, to improve on the single-region inclusion model considered in (Singh, Ding, Ozay, and Yong, 2018) for over-approximating Dubins dynamics, we utilize a tighter/more precise inclusion model in the form of a piecewise affine inclusion model with two subregions that is obtained from the affine abstraction approach in Jin et al. (2020) and is given by:

$$\begin{aligned} A_1 &= [(1, 3) : 0.9048; (2, 4) : 21.781]_{6 \times 6}; \\ A_2 &= [(1, 3) : 0.9048; (2, 4) : 26.6212]_{6 \times 6}; \\ B_1^u &= B_2^u = B_1^w = B_2^w = B = [(3, 1) : 1; (4, 2) : 1; (6, 3) : 1]_{6 \times 3}; \\ F_1 &= [\underline{f}_1, \bar{f}_1] = \begin{bmatrix} (1, 1) : [-3.4680, 5.8193]; \\ (1, 2) : [-9.2068, 9.2068] \end{bmatrix}_{6 \times 1}; \\ F_2 &= [\underline{f}_2, \bar{f}_2] = \begin{bmatrix} (1, 1) : [-4.1616, 6.9832]; \\ (1, 2) : [-10.8352, 10.8352] \end{bmatrix}_{6 \times 1}; \end{aligned}$$

where we used a sparse matrix notation with the size indicated in the subscript. Note that the subscripts $i \in \{1, 2\}$ in the above matrices and vectors denote the indices for the subregions and when using the affine abstraction approach in Jin et al. (2020), the obtained results are such that $\bar{A}_i = \underline{A}_i = A_i$, $\bar{B}_i^u = \underline{B}_i^u = B_i^u$ and $\bar{B}_i^w = \underline{B}_i^w = B_i^w$ for all $i \in \{1, 2\}$.

Combining the abstraction with the intention models and using Euler method for time discretization with sampling time $\delta t = 0.4s$, $\dagger \in \{1, 2\}$, we have the following models:

Inattentive Driver ($\ell = I$) :

$$\begin{aligned} \underline{A}_{I,\dagger} &= \bar{A}_{I,\dagger} = \mathbb{I} + \delta t A_\dagger, \\ \underline{B}_{I,\dagger}^u &= \bar{B}_{I,\dagger}^u = \underline{B}_{I,\dagger}^w = \bar{B}_{I,\dagger}^w = \mathbb{I} + \delta t B, \\ \underline{C}_I &= \bar{C}_I = [(1, 6) : 1]_{1 \times 6}, \underline{D}_I = \bar{D}_I = 0, \underline{V}_I = \bar{V}_I = 1, \\ \underline{f}_{I,\dagger} &= \delta t \underline{f}_\dagger, \bar{f}_{I,\dagger} = \delta t \bar{f}_\dagger, S_{I,\dagger}^x = [(1, 3) : 1; (2, 3) : -1]_{2 \times 6}, \\ \beta_{I,1} &= [25; -20], \beta_{I,2} = [30; -25]; \end{aligned}$$

Cautious Driver ($\ell = C$) :

$$\begin{aligned} \bar{A}_C &= [(6, 3) : -K_{d,C}; (6, 4) : L_{p,C}; (6, 6) : K_{d,C}]_{6 \times 6}, \\ \bar{B}_C &= [(6, 2) : L_{d,C}]_{6 \times 3}, \underline{A}_{C,\dagger} = \bar{A}_{C,\dagger} = \mathbb{I} + \delta t (A_\dagger + \bar{A}_C), \\ \underline{B}_{C,\dagger}^u &= \bar{B}_{C,\dagger}^u = \underline{B}_{C,\dagger}^w = \bar{B}_{C,\dagger}^w = \mathbb{I} + \delta t (B + \bar{B}_C), \\ \underline{C}_C &= \bar{C}_C = [(1, 6) : 1]_{1 \times 6}, \underline{D}_C = \bar{D}_C = 0, \underline{V}_C = \bar{V}_C = 1, \\ \underline{f}_{C,\dagger} &= \delta t \underline{f}_\dagger, \bar{f}_{C,\dagger} = \delta t \bar{f}_\dagger, S_{C,\dagger}^x = [(1, 3) : 1; (2, 3) : -1]_{2 \times 6}, \\ \beta_{C,1} &= [25; -20], \beta_{C,2} = [30; -25]; \end{aligned}$$

Malicious Driver ($\ell = M$) :

$$\begin{aligned} \bar{A}_M &= [(6, 3) : -K_{d,M}; (6, 4) : L_{p,M}; (6, 6) : K_{d,M}]_{6 \times 6}, \\ \bar{B}_M &= [(6, 2) : L_{d,M}]_{6 \times 3}, \underline{A}_{M,\dagger} = \bar{A}_{M,\dagger} = \mathbb{I} + \delta t (A_\dagger - \bar{A}_M), \\ \underline{B}_{M,\dagger}^u &= \bar{B}_{M,\dagger}^u = \underline{B}_{M,\dagger}^w = \bar{B}_{M,\dagger}^w = \mathbb{I} + \delta t (B - \bar{B}_M), \\ \underline{C}_M &= \bar{C}_M = [(1, 6) : 1]_{1 \times 6}, \underline{D}_M = \bar{D}_M = 0, \underline{V}_M = \bar{V}_M = 1, \\ \underline{f}_{M,\dagger} &= \delta t \underline{f}_\dagger, \bar{f}_{M,\dagger} = \delta t \bar{f}_\dagger, S_{M,\dagger}^x = [(1, 3) : 1; (2, 3) : -1]_{2 \times 6}, \\ \beta_{M,1} &= [25; -20], \beta_{M,2} = [30; -25], \end{aligned}$$

where $K_{d,C} = 1$, $K_{d,M} = 0.9$, $L_{p,C} = L_{p,M} = 12$, $L_{d,C} = L_{d,M} = 14$. Further, the initial conditions are:

$$v_e(0) \in [26, 28] \frac{m}{s}, v_o(0) \in [26, 28] \frac{m}{s}, y_e(0) \in [12, 14] m.$$

The lateral position of the ego car is constrained to be between 0 and 26 m, while the other initial conditions, state constraints, input constraints and noise bounds are as in (Singh, Ding, Ozay, and Yong, 2018).

With $J(\cdot) = \|\cdot\|$ and $\lambda = 100$, our proposed AMD algorithm obtained the following inputs: $\{u_1(k)\}_{k=0}^2 = \{-7.5100, -7.4700, 0.0100\}$, $\{u_2(k)\}_{k=0}^2 = \{-0.2065, 0.3568, 0.4400\}$. When compared with the standard AMD (with strict separation constraints), the computational cost of the proposed AMD was around 504 seconds (averaged over 10 runs), while the standard AMD took over 2 days. Furthermore, we compared our AMD algorithm with Niu et al. (2022) using the same settings as the latter, and found that even though the resulting optimal objective value is larger with the proposed AMD algorithm (i.e., it is slightly suboptimal), the computational cost reduces significantly. The proposed AMD algorithm took several minutes to solve, while the method in Niu et al. (2022) took over 5 days. This improvement in computational time is as expected since (multi-)parametric mixed-integer optimization problems² are a much harder class of optimization problems to solve than (regular) mixed-integer optimization problems.

6. CONCLUSION

In this paper, we proposed a novel AMD problem for piecewise affine inclusion systems with bounded uncertainties. Having observed that the integer variables/constraints stemming from the mapping of piecewise inclusion and subregions cause the problem to be very hard to solve with existing optimization solvers, we proposed an alternative bilevel optimization reformulation that moves the integer variables/constraints in the lower level/inner problem to the higher level/outer problem. By leveraging KKT conditions for the lower level that is now free of binary variables, we recast the problem as an equivalent single-level MILP that off-the-shelf solvers can readily solve. Moreover, we introduced a slight extension with a binary variable to

² Interested readers are referred to Pistikopoulos et al. (2020) and references therein for more details, including the computational complexity, of such problems.

indicate the discrimination of each model pair that enables the separation of as many model pairs as possible, if all models cannot be strictly separated. In our future work, we will investigate an extension of our proposed machinery to discriminate among models with signal temporal logic specifications and objective/utility functions.

REFERENCES

- Blackmore, L., Rajamanoharan, S., and Williams, B.C. (2008). Active estimation for jump Markov linear systems. *IEEE Transactions on Automatic Control*, 53(10), 2223–2236.
- Cheong, S. and Manchester, I. (2015). Input design for discrimination between classes of LTI models. *Automatica*, 53, 103–110.
- 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 International Conference on Cyber-Physical Systems*.
- Gurobi (2015). Gurobi optimizer reference manual. URL <http://www.gurobi.com>.
- Harirchi, F., Luo, Z., and Ozay, N. (2016). Model (in)validation and fault detection for systems with polynomial state-space models. In *American Control Conference (ACC)*, 1017–1023.
- Jin, Z., Shen, Q., and Yong, S.Z. (2020). Mesh-based piecewise affine abstraction with polytopic partitions for nonlinear systems. *IEEE Control Systems Letters*, 5(5), 1543–1548.
- Kerestecioglu, F. and Çetin, (2004). Optimal input design for the detection of changes towards unknown hypotheses. *International Journal of Systems Science*, 35(7), 435–444.
- Lou, H. and Si, P. (2009). The distinguishability of linear control systems. *Nonlinear Analysis: Hybrid System*, 3(1), 21–39.
- Marseglia, G.R. and Raimondo, D.M. (2017). Active fault diagnosis: A multi-parametric approach. *Automatica*, 79, 223–230.
- 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.
- Niu, R., Hassaan, S.M., and Yong, S.Z. (2022). A multi-parametric method for active model discrimination of nonlinear systems with temporal logic-constrained switching. In *American Control Conference*, 1652–1658.
- Niu, R., Shen, Q., and Yong, S.Z. (2019). Partition-based parametric active model discrimination design with applications to driver intention estimation. In *European Control Conference*, 3880–3885.
- Patton, R.J., Frank, P.M., and N., C.R. (2000). *Issues of Fault Diagnosis for Dynamic Systems*. Springer London.
- Pistikopoulos, E.N., Diangelakis, N.A., and Oberdieck, R. (2020). *Multi-parametric Optimization and Control*. John Wiley & Sons.
- Raimondo, D.M., Braatz, R.D., and Scott, J.K. (2013). Active fault diagnosis using moving horizon input design. In *European Control Conference*, 3131–3136.
- Raimondo, D.M., Marseglia, G.R., Braatz, R.D., and Scott, J.K. (2016). Closed-loop input design for guaranteed fault diagnosis using set-valued observers. *Automatica*, 74, 107–117.
- Scott, J.K., Findeison, R., Braatz, R.D., and Raimondo, D.M. (2014). Input design for guaranteed fault diagnosis using zonotopes. *Automatica*, 50(6), 1580–1589.
- Shen, Q., Niu, R., and Yong, S.Z. (2022). Tractable model discrimination for safety-critical systems with disjunctive and coupled constraints. *Nonlinear Analysis: Hybrid Systems*, 46, 101217.
- Singh, K., Ding, Y., Ozay, N., and Yong, S.Z. (2018). Input design for nonlinear model discrimination via affine abstraction. *IFAC PapersOnLine*, 51(16), 175–180.
- Singh, K., Shen, Q., and Yong, S.Z. (2018). Mesh-based affine abstraction of nonlinear systems with tighter bounds. In *IEEE Conference on Decision and Control*, 3056–3061.
- Tabatabaeipour, S. (2015). Active fault detection and isolation of discrete-time linear time-varying systems: a set-membership approach. *International Journal of Systems Science*, 46(11), 1917–1933.
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S. (2003). A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. *Computers & Chemical Eng.*, 27(3), 293–311.

APPENDIX

$$\begin{aligned}
 M_\ell &= \vec{\text{vec}}_{i=1}^{q_\ell} \{M_{\ell,i}\}, \Omega_\ell^x = \vec{\text{vec}}_{i=1}^{q_\ell} \{\Omega_{\ell,i}^x\}, F_\ell = \vec{\text{vec}}_{i=1}^{q_\ell} \{F_{\ell,i,T}\}, \\
 \beta_\ell &= \vec{\text{vec}}_{i=1}^{q_\ell} \{\beta_{\ell,i,T}\}, \alpha_\ell = \vec{\text{vec}}_{i=1}^{q_\ell} \{\alpha_{\ell,i}\}, \zeta_\ell = \vec{\text{vec}}_{i=1}^{q_\ell} \{\zeta_{\ell,i}\}, \\
 M^\ell &= \begin{bmatrix} \text{diag} \{M_\ell\} \\ \ell=\{\dagger,\dagger\} \\ \text{diag} \{\Omega_\ell\} \\ \ell=\{\dagger,\dagger\} \end{bmatrix}, F_T^\ell = \begin{bmatrix} \text{diag} \{F_\ell\} + \text{diag} \{\alpha_\ell\} \\ \ell=\{\dagger,\dagger\} \\ \text{diag} \{\beta_\ell\} + \text{diag} \{\zeta_\ell\} \\ \ell=\{\dagger,\dagger\} \end{bmatrix}, \\
 \Omega_{\ell,i}^x &= [\text{diag}_T \{S_{\ell,i}^x\} \mathbf{0}], \\
 E^\ell &= \text{diag} \{E_\ell\}, G^\ell = \text{vec}_{\ell=\{\dagger,\dagger\}} \{G_\ell\}, E_\ell = \text{diag}_T \{C_\ell\}, G_\ell = \text{vec}_T \{g_\ell\}, \\
 \beta_{\ell,i,T} &= \text{vec}_{T-1} \{\beta_{\ell,i}\}, R_1^\ell = [M^\ell \Gamma_d^\ell \Gamma_w^\ell \mathbf{0}], \\
 R_2^\ell &= \begin{bmatrix} E_i & -E_j & F_{d,i} & -F_{d,j} & \mathbf{0} & \mathbf{0} & F_{v,i} & -F_{v,j} \\ -E_i & E_j & -F_{d,i} & F_{d,j} & \mathbf{0} & \mathbf{0} & -F_{v,i} & F_{v,j} \end{bmatrix}, \\
 r_1^\ell &= F_T^\ell, S_1^\ell = \Gamma_u^\ell, r_2^\ell = \begin{bmatrix} G_i - G_j \\ G_j - G_i \end{bmatrix}, S_2^\ell = \begin{bmatrix} F_{u,i} - F_{u,j} \\ F_{u,j} - F_{u,i} \end{bmatrix}, \\
 M_{\ell,i} &= \begin{bmatrix} \underline{A}_{\ell,i} & -\mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ -\bar{A}_{\ell,i} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \underline{A}_{\ell,i} & -\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & -\bar{A}_{\ell,i} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \underline{A}_{\ell,i} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & -\bar{A}_{\ell,i} & \mathbf{I} \end{bmatrix}_{2nT \times 2n(T+1)}, \\
 F_{\ell,i,T} &= \begin{bmatrix} f_{\ell,i}^\top & -\bar{f}_{\ell,i}^\top & \underline{f}_{\ell,i}^\top & -\bar{f}_{\ell,i}^\top & \cdots & f_{\ell,i}^\top & -\bar{f}_{\ell,i}^\top \end{bmatrix}_{2nT \times 1}^\top, \\
 \text{For } \star &= \{w, u\}: \Gamma_{\ell,i}^\star = \text{diag}_T \left\{ \begin{bmatrix} \underline{B}_{\ell,i}^\star \\ -\bar{B}_{\ell,i}^\star \end{bmatrix} \right\}, F_\ell^\star = \text{diag}_T \{D_{\star,\ell}\}, \\
 \Omega_{\ell,i}^\star &= \text{diag}_T \{S_{\ell,i}^\star\}, \Gamma_\star^\ell = \begin{bmatrix} \text{diag} \text{vec}_{i=1}^{q_\ell} \{\Gamma_{\ell,i}^\star\} \\ \ell=\{\dagger,\dagger\} \\ \text{diag} \text{vec}_{i=1}^{q_\ell} \{\Omega_{\ell,i}^\star\} \\ \ell=\{\dagger,\dagger\} \end{bmatrix}, F_\star^\ell = \text{vec}_{\ell=\{\dagger,\dagger\}} \{F_\ell^\star\}, \\
 \Omega_\ell^\star &= \text{vec}_{i=1}^{q_\ell} \{\Omega_{\ell,i}^\star\}, \Gamma_\ell^\star = \text{vec}_{i=1}^{q_\ell} \{\Gamma_{\ell,i}^\star\}, \\
 \bar{P}^\ell &= [M_\ell \Gamma_d^\ell \Gamma_w^\ell], \bar{P}_u^\ell = [-\Gamma_u^\ell], \bar{p}^\ell = -[F_\ell + \alpha_\ell], \\
 \bar{\bar{P}}_x^\ell &= \begin{bmatrix} \Omega_\ell^x & \Omega_\ell^d & \Omega_\ell^w \\ [\bar{P}_{x,\ell} \mathbf{0}] & \mathbf{0} & \mathbf{0} \end{bmatrix}, \bar{\bar{P}}_u^\ell = -\begin{bmatrix} \Omega_u^x \\ \mathbf{0} \end{bmatrix}, \bar{\bar{p}}^\ell = -\begin{bmatrix} \beta_\ell + \zeta_\ell \\ -\bar{p}_{x,\ell} \end{bmatrix}, \\
 \bar{\bar{H}}^\ell &= [\text{diag} \{\bar{P}_0^\ell, Q_{d,\ell}, Q_{w,\ell}\}], \bar{\bar{h}}^\ell = [\bar{p}_0^\ell \bar{p}_d^\ell \bar{q}_w^\ell]^\top.
 \end{aligned}$$