2022 European Control Conference (ECC)
July 12-15, 2022. London, United Kingdom

Multi-Model Affine Abstraction of Nonlinear Systems with Model Discrimination Guarantees

Syed M. Hassaan, Zeyuan Jin and Sze Zheng Yong

Abstract— **This paper presents a novel optimization-based method for multi-model affine abstraction (i.e., for simultaneous model reduction of multiple models), which solves for the existence of affine abstractions of a pair of different nonlinear systems with guarantees of model discrimination with the minimum detection time T under worst-case uncertainties and approximation errors. Our approach combines mesh-based affine abstraction methods with T-distinguishability analysis in the literature into a bilevel bilinear optimization problem. Then, to obtain a tractable solution, we leverage robust optimization techniques and a suitable change of variables to obtain a sufficient linear program (LP). Finally, the efficacy of proposed methods is illustrated by several numerical examples.**

I. Introduction

Abstraction-based methods (i.e., a specific class of model reduction techniques) can simplify the process for analyzing and controlling smart and complex (nonlinear or hybrid) systems [1] by computing a simpler but over-approximated system, i.e., to find an over-approximation of its original dynamics $f(\cdot)$ by a simpler inclusion model $\overline{f}(\cdot)$ and $\underline{f}(\cdot)$ as bracketing functions or framers, such that for all x in defined domain, $\underline{f}(x) \leq f(x) \leq \overline{f}(x)$. By design, abstraction models must include all possible behaviors of the original system to maintain properties of interest. For instance, to verify that a given complex system satisfies certain properties, we can test for the desired property on the abstracted simple system which provides an equivalent or sufficient result as when testing for the property on the original complex system.

Literature Review. Numerous abstraction approaches have been developed for different classes of systems in the literature, e.g., nonlinear systems [2]–[5], hybrid systems [6], and uncertain affine and nonlinear systems [7], [8]. Two important classes of abstraction methods are *symbolic* approaches [1], [9] and *hybridization* [2], [10]. In these methods, the growth of the number of symbolic states and inputs (or partitions) is exponential with state and input dimensions. Thus, [11] presented a incremental method to find abstraction by solving a sequence of linear programs (LP) to overcome the scalability issue. In addition, when the exact dynamic model is unknown, the abstraction can be identified using data-driven approaches [12], [13].

Another specific interest of this paper is the model discrimination/identification problem, which seeks to distinguish between models based on the compatibility and consistency of newly observed input-output data with the models using a model (in)validation framework [14], [15]. Multiple methods to solve the model invalidation problem has been proposed for various types of systems, e.g., linear parameter varying systems [16], [17], nonlinear systems [18], uncertain systems [8], switched auto-regressive models [19] and switched affine systems [15], [20]. Further, given a pair of models, the notion of T-detectability/distinguishability is introduced in [15], [20] to analyze the distinction of these two models, which is defined as the upper bound on the required time horizon T to distinguish the model pair, if such a T exists. The notion of T-distinguishability is also related to state/mode distinguishability in the literature for switched linear [21], finite-state [22] and switched nonlinear [23] systems.

Contributions. In this paper, we propose a mesh-based method to compute affine abstractions for a pair of different nonlinear systems and to find the minimum guaranteed detection time T such that the two abstractions with the same sets of initial state and noise signals are guaranteed to be distinguished within finite time steps T under worst-case realizations of uncertainties and abstraction/modeling errors. In contrast to existing two-step methods that separate the processes of finding abstraction and analyzing distinguishability, e.g., [24], we propose a bilevel optimization problem that, if solved, simultaneously provide the abstraction model and required time horizon T to distinguish one model from the other. To solve the bilevel problem, we rely on robust optimization techniques [25] to obtain a single-level dual problem. However, this problem involves bilinear terms that might lead to intractability; thus, we further convert the bilinear optimization problem into a sufficient linear programming (LP) problem by a suitable change of variables. To the best of our knowledge, this is the first work that can simultaneously find abstractions for multiple different nonlinear systems to maximize the "difference" between the abstractions to facilitate model discrimination, e.g., in fault detection and (intent) model estimation applications. Finally, we demonstrate the effectiveness of our approaches to distinguish pairs of 1D and 2D models.

II. Preliminaries

Notation. We use $\mathbb{R}^n$ to denote a real vector of length n, $\mathbb{Z}_a^b$ to denote a set with integers from a to b, I_n to denote an n-by-n identity matrix and $\mathbb{1} \in \mathbb{R}^n$ to denote a vector of ones, i.e., $\mathbb{1} = [1, 1, \ldots, 1]^\top$. Moreover, we use $\underset{i}{vec}(D_i)$ to denote a row vector constructed by concatenating scalars/vectors D_i. For a vector v and a matrix M, $|v|$, $\|v\|_i$ and $\|M\|_i$ denote their component-wise absolute values and their (induced) i-norms with $i = \{1, 2, \infty\}$.

The authors are with the School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, USA (email: {shassaan,zjin43,szyong}@asu.edu). This work was supported in part by DARPA grant D18AP00073 and NSF grant CNS-1943545.

A. Modeling Framework

In this paper, we consider a pair of nonlinear discrete-time system dynamics over a common domain given by:

$$x_{i,k+1} = f_i(x_{i,k}, u_{i,k}) + w_{i,k}, \quad (1)$$
$$y_{i,k} = C_i x_{i,k} + v_{i,k}, \quad (2)$$

where the subscript $i \in \{1,2\}$ is used to refer to the different models and $x_{i,k} \in \mathcal{X} \subseteq \mathbb{R}^n$ is the system state, $w_{i,k} \in \mathcal{W} \subseteq \mathbb{R}^n$ is measurement noise linearly affecting the dynamics and $u_{i,k} \in \mathcal{U} \subseteq \mathbb{R}^m$ is the known control input with a bounded and closed interval domain $\mathcal{U}$. $y_{i,k} \in \mathcal{Y} \subseteq \mathbb{R}^m$ is the observation and $v_{i,k} \in \mathcal{V} \subseteq \mathbb{R}^m$ is the measurement noise. We assume $\mathcal{X}$, $\mathcal{U}$ and $\mathcal{W}$ to be known bounded and closed interval domains, with specifically $\|w_{i,k}\|_\infty \leq \eta_w$ and $\|v_{i,k}\|_\infty \leq \eta_v$. The vector fields $f_i : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$ are continuous functions. Moreover, $(x_s, u_s) \in \mathbb{R}^{n+m}$ represents a *sample point*, and a partition of the domain is denoted by $I \triangleq \mathcal{X} \times \mathcal{U}$. The following definitions are introduced to define a sample set $\mathcal{M} \subset \mathcal{X} \times \mathcal{U}$:

Definition 1 (Uniform Mesh, Mesh Elements, Grid Points and Diameter [11])**.** *A uniform mesh of each domain I is a collection of polytopic partitions, called mesh elements, with a total of $s_{\max}$ number of points, called grid points, uniformly distributed along all directions and dimensions. The diameter δ of a uniform mesh is the greatest distance between vertices of each mesh element.*

Definition 2 (Sample and Vertex Sets)**.** *A sample set is the set of grid points, denoted as $\mathcal{M}$ and by construction, the convex hull of $\mathcal{M}$ is the domain I, i.e., $I = Conv(\mathcal{M})$. The set of all vertices of $\mathcal{X}$ is called the vertex set of state x, and denoted as $\mathcal{X}_c$. Similarly, the set of all vertices of $\mathcal{U}$ is called the vertex set of input u, and denoted as $\mathcal{U}_c$. Because we are considering both $\mathcal{X}$ and $\mathcal{U}$ to be bounded, they both are polytopes and have well-defined vertex sets.*

For ease of exposition and to reduce notational burden, we introduce the following assumption on the model outputs:

A1. *For the given nonlinear dynamics* (1)*, each system has full-state output with measurement noise, i.e., $C = I$.*

Note that our proposed algorithm can be easily extended to consider any linear matrix C and/or more than two models.

B. Abstraction/Over-approximation

The abstraction model for f_i for each $i \in \{1,2\}$ is defined as a pair of functions $\underline{f}_i$ and $\overline{f}_i$ which over-approximate/frame the original function $f_i(\cdot) : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$ such that $\forall (x_s, u_s) \in \mathcal{M}$, $\underline{f}_i(x_s, u_s) \leq f_i(x_s, u_s) \leq \overline{f}_i(x_s, u_s)$. Specifically, we aim to find an affine abstraction model $\mathcal{H}_i$ for each $i \in \{1,2\}$ with respect to (1) over domain $I \triangleq \mathcal{X} \times \mathcal{U}$ which is given by:

$$\begin{aligned} \underline{A}_{i,k} x_{i,k} + \underline{B}_{i,k} u_{i,k} + \underline{h}_{i,k} + w_{i,k} &\leq x_{i,k+1} \\ &\leq \overline{A}_{i,k} x_{i,k} + \overline{B}_{i,k} u_{i,k} + \overline{h}_{i,k} + w_{i,k}, \end{aligned} \quad (3)$$

where $\underline{A}_{i,k}, \underline{B}_{i,k}, \underline{h}_{i,k}, \overline{A}_{i,k}, \overline{B}_{i,k}$ and $\overline{h}_{i,k}$ are (time-varying) matrices/vectors of appropriate dimensions.

Remark 1. *If desired, a global affine abstraction model $\mathcal{H}_i^g$ can be found by further abstracting* (3) *in a post-processing step (with given $\underline{A}_{i,k}, \underline{B}_{i,k}, \underline{h}_{i,k}, \overline{A}_{i,k}, \overline{B}_{i,k}$ and $\overline{h}_{i,k}$ for $k \in \mathbb{Z}_0^{T-1}$ and $i \in \{1,2\}$) such that: $\underline{A}_i x_{i,k} + \underline{B}_i u_{i,k} + \underline{h}_i + w_{i,k} \leq \underline{A}_{i,k} x_{i,k} + \underline{B}_{i,k} u_{i,k} + \underline{h}_{i,k} + w_{i,k} \leq x_{i,k+1} \leq \overline{A}_{i,k} x_{i,k} + \overline{B}_{i,k} u_{i,k} + \overline{h}_{i,k} + w_{i,k} \leq \overline{A}_i x_{i,k} + \overline{B}_i u_{i,k} + \overline{h}_i + w_{i,k}$, where $\underline{A}_i, \underline{B}_i, \underline{h}_i, \overline{A}_i, \overline{B}_i$ and $\overline{h}_i$ are time-invariant matrices/vectors of appropriate dimensions. Our future work will consider an extension for directly finding global abstraction models $\mathcal{H}_i^g$.*

In preparation for finding the affine abstraction for f_i (performed dimension-wise, i.e., for $f_i^{(j)}, \forall j \in \mathbb{Z}_1^n$), we rely on the following definitions and proposition:

Definition 3 (Abstraction Error [3])**.** *The abstraction error of an affine abstraction model $\mathcal{H}_i$ of a sampled nonlinear function $f_i(x_s, u_s)$ over its domain I is defined as $\theta_i = \max_{(x_s,u_s) \in \mathcal{M}} \|\overline{f}_i(x_s, u_s) - \underline{f}_i(x_s, u_s)\|_\infty$.*

Proposition 1 ([26, Theorem 4.1 & Lemma 4.3])**.** *Let S be an $(n+m)$-dimensional mesh element such that $S \subseteq \mathbb{R}^{n+m}$ with diameter δ (see Definition 1). Let $f_i^{(j)} : S \to \mathbb{R}$ be a nonlinear function and let $f_{l,i}^{(j)}$ be the linear interpolation of $f_i^{(j)}(\cdot)$ evaluated at the vertices of the mesh element S. Then, the approximation error σ defined as the maximum error between $f_i^{(j)}$ and $f_{l,i}^{(j)}$ on S, i.e., $\sigma = \max_{s \in S}(|f_i^{(j)}(s) - f_{l,i}^{(j)}(s)|)$, is upper-bounded by*

(i) $\sigma \leq 2 \max_{s \in S} \|f_i^{(j)}(s)\|_\infty$, if $f_i^{(j)} \in C^0$ on S,
(ii) $\sigma \leq \lambda \delta_s$, if $f_i^{(j)}$ is Lipschitz continuous on S,
(iii) $\sigma \leq \delta_s \max_{s \in S} \|{f'}_i^{(j)}(s)\|_2$, if $f_i^{(j)} \in C^1$ on S,
(iv) $\sigma \leq \frac{1}{2}\delta_s^2 \max_{s \in S} \|{f''}_i^{(j)}(s)\|_2$, if $f_i^{(j)} \in C^2$ on S,

where C^0, C^1 and C^2 are sets of continuous, continuously differentiable and twice continuously differentiable functions respectively, λ is the Lipschitz constant, ${f'}_i^{(j)}(s)$ is the Jacobian of $f_i^{(j)}(s)$, ${f''}_i^{(j)}(s)$ is the Hessian of $f_i^{(j)}(s)$ and δ_s satisfies $\delta_s \leq \sqrt{\frac{n+m}{2(n+m+1)}}\delta$.

Since this paper involves a pair of nonlinear functions f_1 and f_2, we denote their respective approximation errors with σ_1 and σ_2 for the remainder of the paper.

C. Length-T Behavior

To solve the model discrimination problem, we utilize the definition in [15] of the length-T behaviors of the original nonlinear model $\mathcal{G}$ and the abstracted model $\mathcal{H}$:

Definition 4 (Length-T Behaviors of Original and Abstracted Models $\mathcal{G}_i$ and $\mathcal{H}_i$)**.** *The length-T behaviors of the original (nonlinear) model $\mathcal{G}_i$ and affine abstracted model $\mathcal{H}_i$ with the model output governed by Assumption 1 are the sets of all length-T output trajectories compatible with $\mathcal{G}_i$ and $\mathcal{H}_i$, respectively, given by the sets:*

$$\mathcal{B}^T(\mathcal{G}_i) := \{\{u_k, y_k\}_{k=0}^{T-1} \mid \exists x_k \in \mathcal{X}, u_k \in \mathcal{U}, y_k \in \mathcal{Y}, w_k \in \mathcal{W}, \quad v_k \in \mathcal{V}, \quad \textit{for } k \in \mathbb{Z}_{T-1}^0, \quad \textit{s.t. } (1)\text{–}(2) \textit{ hold}\},$$
$$\mathcal{B}^T(\mathcal{H}_i) := \{\{u_k, y_k\}_{k=0}^{T-1} \mid \exists x_k \in \mathcal{X}, u_k \in \mathcal{U}, y_k \in \mathcal{Y}, w_k \in \mathcal{W}, \quad v_k \in \mathcal{V}, \quad \textit{for } k \in \mathbb{Z}_{T-1}^0, \quad \textit{s.t. } (2)\text{–}(3) \textit{ hold}\}.$$

Thus, by definition, $\mathcal{B}^T(\mathcal{G}_i) \subseteq \mathcal{B}^T(\mathcal{H}_i)$.

III. Problem Formulation

When two nonlinear models, usually encompassing different behaviors of a system over the same domain, are given, the problem that we aim to tackle in this paper is to find an affine over-approximation of each of these models such that there is some way to discriminate between these abstractions when utilized. The discrimination solution using the abstraction models is sufficient for the problem with the original nonlinear models because $\mathcal{B}^T(\mathcal{G}_i) \subseteq \mathcal{B}^T(\mathcal{H}_i)$ (cf. Definition 4). For that purpose, we look into the concept of T-distinguishability for model discrimination. Formally, we define the problem we consider in this paper as:

Problem 1. *For a given pair of nonlinear dynamics defined in* (1) *with* $(x,u) \in \mathcal{X} \times \mathcal{U}$*, find affine abstractions* $\mathcal{H}_i$*,* $i \in \{1,2\}$ *of the form* (3) *for their respective* $f_i(x,u)$ *over* I*, such that both* $\mathcal{H}_i$*'s are T-distinguishable for a minimum* T*, i.e., find the minimum horizon* T *and* $\mathcal{H}_i$ *such that:*

$$\bigcap_{i=1}^{2} \mathcal{B}^T(\mathcal{H}_i) = \emptyset. \tag{4}$$

To solve Problem 1, we will also consider a special case of the problem, where we find the abstractions that are T-distinguishable for a specified T.

Problem 1.1. *Given a pair of nonlinear dynamics defined in* (1) *with* $(x,u) \in \mathcal{X} \times \mathcal{U}$*, determine whether there exist affine abstractions* $\mathcal{H}_i$*,* $i \in \{1,2\}$ *of the form* (3) *for their respective* $f_i(x,u)$ *over* I*, such that both* $\mathcal{H}_i$*'s are T-distinguishable for a specified* T*, i.e., for a fixed* T*, determine if there exist* $\mathcal{H}_i$*,* $i \in \{1,2\}$*, such that:*

$$\bigcap_{i=1}^{2} \mathcal{B}^T(\mathcal{H}_i) = \emptyset. \tag{5}$$

IV. Main Results

In this section, we propose solution approaches that can be used to solve Problem 1. Specifically, we start by leveraging the concept of affine abstraction of the nonlinear functions from [3], and combine it with the T-distinguishability conditions for the abstracted models from [15]. The purpose of applying the T-distinguishability condition is to aid in yielding affine abstractions of the pair of functions that can be distinguished from each other within T time steps, hence *separating* the two models, solving the second part of Problem 1.

To design such an optimization problem, we first start by fixing the horizon to an arbitrary T. Then, for the two nonlinear dynamics (1), we first propose an optimization problem to tackle the Problem 1.1 as follows:

Proposition 2. *Given a pair of nonlinear systems of the form* (1) *with initial state and input values bounded by* $x_0 \in \mathcal{X}_0 \subseteq \mathcal{X}$ *and* $u_0 \in \mathcal{U}_0 \subseteq \mathcal{U}$*, respectively, along with Assumption 1, there exist time-varying affine abstractions* $\mathcal{F}_{i,k} = \{\overline{f}_{i,k}, \underline{f}_{i,k}\}$ *for each system* $i = \{1,2\}$ *that are* T*-distinguishable for a known* T *if the following optimization problem is feasible:*

$$\min_{\substack{\overline{A}_{1,k},\underline{A}_{1,k},\overline{B}_{1,k},\underline{B}_{1,k},\overline{h}_{1,k},\underline{h}_{1,k},\\ \overline{A}_{2,k},\underline{A}_{2,k},\overline{B}_{2,k},\underline{B}_{2,k},\overline{h}_{2,k},\underline{h}_{2,k},\\ \theta_1,\theta_2,\epsilon}} -\epsilon + d(\theta_1+\theta_2)$$

subject to:

$$\begin{aligned}
&\forall((x_s,u_s) \in \mathcal{M}, x_c \in \mathcal{X}_c, u_c \in \mathcal{U}_c, i \in \{1,2\}, k \in \mathbb{Z}_0^{T-1}):\\
&\quad \overline{A}_{i,k}x_s + \overline{B}_{i,k}u_s + \overline{h}_{i,k} \geq f_i(x_s,u_s),\\
&\quad \underline{A}_{i,k}x_s + \underline{B}_{i,k}u_s + \underline{h}_{i,k} \leq f_i(x_s,u_s),\\
&(\overline{A}_{i,k}-\underline{A}_{i,k})x_c + (\overline{B}_{i,k}-\underline{B}_{i,k})u_c + \overline{h}_{i,k} - \underline{h}_{i,k} \leq \theta_i \mathbb{1},
\end{aligned} \tag{6a}$$

$$\begin{aligned}
&\gamma^\star \triangleq \min_{x,u,y,z,\gamma} \gamma \geq \epsilon\\
&\text{subj. to: } \forall k \in \mathbb{Z}_0^{T-1}:\\
&x_{k+1} \leq \overline{A}_{1,k}x_k + \overline{B}_{1,k}u_k + \overline{h}_{1,k} + \sigma_1 + \eta_w \mathbb{1},\\
&x_{k+1} \geq \underline{A}_{1,k}x_k + \underline{B}_{1,k}u_k + \underline{h}_{1,k} - \sigma_1 - \eta_w \mathbb{1},\\
&y_{k+1} \leq \overline{A}_{2,k}y_k + \overline{B}_{2,k}z_k + \overline{h}_{2,k} + \sigma_2 + \eta_w \mathbb{1},\\
&y_{k+1} \geq \underline{A}_{2,k}y_k + \underline{B}_{2,k}z_k + \underline{h}_{2,k} - \sigma_2 - \eta_w \mathbb{1},\\
&P_k x_k \leq p_k, P_k y_k \leq p_k, Q_k u_k \leq q_k, Q_k z_k \leq q_k,\\
&|x_k - y_k| \leq (\gamma + 2\eta_v)\mathbb{1},
\end{aligned} \tag{6b}$$

where d *is an arbitrarily small scalar,* x_s *and* u_s *are values at the grid points on the mesh* $\mathcal{M}$*, and* x_c *and* u_c *are vertices of the domain* $\mathcal{X} \times \mathcal{U}$*. The matrix-vector pairs* (P_k, p_k) *and* (Q_k, q_k) *are polytopic representations of the domains* $\mathcal{X}_0$ *and* $\mathcal{U}_0$ *when* $k = 0$*, and* $\mathcal{X}$ *and* $\mathcal{U}$ *otherwise, respectively.*

Proof. To obtain the abstraction $\mathcal{H}_i$ of f_i, it needs to fulfil the property given in (3). Moreover, for the T-distinguishability condition with Assumption 1, the following needs to hold for all $k \in \mathbb{Z}_0^{T-1}$:

$$\begin{aligned}
&x_{k+1} \leq \overline{A}_{1,k}x_k + \overline{B}_{1,k}u_k + \overline{h}_{1,k} + \sigma_1 + w_k,\\
&x_{k+1} \geq \underline{A}_{1,k}x_k + \underline{B}_{1,k}u_k + \underline{h}_{1,k} - \sigma_1 - w_k,\\
&y_{k+1} \leq \overline{A}_{2,k}y_k + \overline{B}_{2,k}z_k + \overline{h}_{2,k} + \sigma_2 + w_k,\\
&y_{k+1} \geq \underline{A}_{2,k}y_k + \underline{B}_{2,k}z_k + \underline{h}_{2,k} - \sigma_2 - w_k,\\
&P_k x_k \leq p_k, P_k y_k \leq p_k, Q_k u_k \leq q_k, Q_k z_k \leq q_k,\\
&|x_k + v_{1,k} - y_k - v_{2,k}| \leq \gamma \mathbb{1},
\end{aligned}$$

The above can be sufficiently described by replacing the noise variables w and v with their bounds, resulting in:

$$\begin{aligned}
&x_{k+1} \leq \overline{A}_{1,k}x_k + \overline{B}_{1,k}u_k + \overline{h}_{1,k} + \sigma_1 + \eta_w \mathbb{1},\\
&x_{k+1} \geq \underline{A}_{1,k}x_k + \underline{B}_{1,k}u_k + \underline{h}_{1,k} - \sigma_1 - \eta_w \mathbb{1},\\
&y_{k+1} \leq \overline{A}_{2,k}y_k + \overline{B}_{2,k}z_k + \overline{h}_{2,k} + \sigma_2 + \eta_w \mathbb{1},\\
&y_{k+1} \geq \underline{A}_{2,k}y_k + \underline{B}_{2,k}z_k + \underline{h}_{2,k} - \sigma_2 - \eta_w \mathbb{1},\\
&P_k x_k \leq p_k, P_k y_k \leq p_k, Q_k u_k \leq q_k, Q_k z_k \leq q_k,\\
&|x_k - y_k| \leq (\gamma + 2\eta_v)\mathbb{1},
\end{aligned} \tag{7}$$

where the objective is to minimize γ. But at the same time, we want to find abstractions that maximize the minimum γ, giving rise to the condition $\max \min \gamma$, resulting in the bilevel problem:

$$\max \epsilon \text{ subj. to: } \epsilon \leq \min \gamma. \tag{8}$$

Further, since the separating abstraction models may not be tight, a secondary goal is to enforce tightness of $\mathcal{H}_i$ over its corresponding f_i, i.e., its abstraction error (cf. Definition 3) is to be minimized, implying:

$$\min \theta_i \text{ subj. to: } (\overline{f}_i(x_s,u_s) - \underline{f}_i(x_s,u_s)) \leq \theta_i, \quad \forall (x_s,u_s) \in \mathcal{M}. \tag{9}$$

Combining the constraints (3), (9), (7) and considering the weighted sum of the costs in (9) and (8) in the form of $-\epsilon + d(\theta_1+\theta_2)$ with d chosen to be very small since tightness is secondary to separation guarantees, the optimization problem, by construction, solves Problem 1.1 if a solution to (6) exists. □

Observing the optimization problem (6), it can be seen that it is a bilevel problem that also involves bilinear terms in constraints (6b), making it extremely hard to solve. To obtain a tractable problem that can be easily solved using off-the-shelf solvers, we first utilize robustification techniques to obtain an equivalent single-level dual problem as follows:

Lemma 1. *A robust dual of the solution in Proposition 2 takes the following form:*

$$\min_{\substack{\overline{A}_{1,k},\underline{A}_{1,k},\overline{B}_{1,k},\underline{B}_{1,k},\overline{h}_{1,k},\underline{h}_{1,k},\\ \overline{A}_{2,k},\underline{A}_{2,k},\overline{B}_{2,k},\underline{B}_{2,k},\overline{h}_{2,k},\underline{h}_{2,k},\\ \theta_1,\theta_2,\epsilon}} -\epsilon + d(\theta_1+\theta_2)$$

subject to:

$$\forall((x_s,u_s)\in\mathcal{M}, x_c\in\mathcal{X}_c, u_c\in\mathcal{U}_c, i\in\{1,2\}, k\in\mathbb{Z}_0^{T-1}):$$

$$\begin{aligned} &\overline{A}_{i,k}x_s+\overline{B}_{i,k}u_s+\overline{h}_{i,k}\geq f_i(x_s,u_s),\\ &\underline{A}_{i,k}x_s+\underline{B}_{i,k}u_s+\underline{h}_{i,k}\leq f_i(x_s,u_s),\\ &(\overline{A}_{i,k}-\underline{A}_{i,k})x_c+(\overline{B}_{i,k}-\underline{B}_{i,k})u_c+\overline{h}_{i,k}-\underline{h}_{i,k}\leq\theta_i\mathbb{1}, \end{aligned} \tag{10a}$$

$$\begin{aligned} &\Pi^\top\begin{bmatrix} r_1+\sigma+\eta_w\mathbb{1}\\ r_2\\ c\\ g\end{bmatrix}\leq-\epsilon,\\ &\Pi^\top\begin{bmatrix} R_1 & S_1 & 0\\ R_2 & 0 & -\mathbb{1}\\ C & 0 & 0\\ 0 & G & 0\end{bmatrix}=\begin{bmatrix}0 & 0 & 1\end{bmatrix}, \Pi^\top\geq 0, \end{aligned} \tag{10b}$$

where R_1, S_1, R_2, C, G, r_1, σ, r_2, c and g are stacked matrices and vectors defined in the Appendix.

Proof. Since the inner problem (6b) constraints involve a finite time horizon T, we can stack the constraint inequalities over the horizon $k\in\mathbb{Z}_0^{T-1}$ to obtain the following inequalities for $i\in\{1,2\}$:

$$\begin{aligned} \underline{M}_i x+\underline{N}_i u+\underline{F}_i-\bar{\sigma}_i-\eta_w\mathbb{1}\leq 0,\\ \overline{M}_i x+\overline{N}_i u+\overline{F}_i-\bar{\sigma}_i-\eta_w\mathbb{1}\leq 0,\\ |x-y|\leq(\gamma+2\eta_v)\mathbb{1},\\ Px\leq p, Py\leq p, Qu\leq q, Qz\leq q, \end{aligned} \tag{11}$$

where $\underline{M}_i$, $\underline{N}_i$, $\underline{F}_i$, $\overline{M}_i$, $\overline{N}_i$ and $\overline{F}_i$ for $i\in\{1,2\}$ are the time-stacked matrices defined in the Appendix and x, y, u, z and $\bar{\sigma}_i$ are stacked versions of the corresponding signals defined as:

$$\begin{aligned} x=[x_0^\top,\cdots,x_T^\top]^\top, u=[u_0^\top,\cdots,u_{T-1}^\top]^\top,\\ y=[y_0^\top,\cdots,y_T^\top]^\top, z=[z_0^\top,\cdots,z_{T-1}^\top]^\top,\\ \bar{\sigma}_i=[\sigma_i^{(1)},\cdots,\sigma_i^{(T)}]^\top\in\mathbb{R}^{nT}\ \forall i\in\{1,2\}. \end{aligned}$$

Now, as x and y correspond to state trajectories following the abstracted dynamics of f_1 and f_2 respectively, we further stack the equations in (11) to obtain the following:

$$\begin{aligned} R_1\bar{x}+S_1\bar{u}\leq r_1+\sigma+\eta_w\mathbb{1}\triangleq\bar{r}_1,\\ R_2\bar{x}\leq r_2+\gamma\mathbb{1},\ C\bar{x}\leq c,\ G\bar{u}\leq g, \end{aligned} \tag{12}$$

where R_1, S_1, r_1, R_2, r_2, C and c are defined in Appendix and $\bar{x}$, $\bar{u}$ and σ are defined as:

$$\bar{x}=[x^\top,y^\top]^\top, \bar{u}=[u^\top,z^\top]^\top, \sigma=[\bar{\sigma}_1^\top,\bar{\sigma}_2^\top]^\top.$$

Rewriting the problem (6b) using the stacked forms in (12), we obtain:

$$[0\ 0\ -1][\bar{x}^\top\ \bar{u}^\top\ \gamma]^\top\leq-\epsilon,\ \forall[\bar{x}^\top\ \bar{u}^\top\ \gamma]^\top\in J(R_1,S_1,r_1), \tag{13}$$

where

$$J(R_1,S_1,r_1)\triangleq\left\{\begin{bmatrix}\bar{x}\\ \bar{u}\\ \gamma\end{bmatrix}:\begin{bmatrix} R_1 & S_1 & 0\\ R_2 & 0 & -\mathbb{1}\\ C & 0 & 0\\ 0 & G & 0\end{bmatrix}\begin{bmatrix}\bar{x}\\ \bar{u}\\ \gamma\end{bmatrix}\leq\begin{bmatrix}\bar{r}_1\\ r_2\\ c\\ g\end{bmatrix}\right\}. \tag{14}$$

As all the constraints involved in (13)–(14) are linear in the decision variables, applying duality theory from the robust optimization literature, e.g., [25], results in a dual problem given in (10b) with the dual variable Π. ☐

Note that the optimization problem in (10) still involves bilinear terms in (10b). To overcome this, a suitable change of variables is applied to obtain a sufficient problem that is linear in the decision variables involved. The resulting optimization problem is as follows:

Theorem 1. *Given a pair of systems as nonlinear functions of the form* (1) *with initial state bounded by $x_0\in\mathcal{X}_0\subseteq\mathcal{X}$ and initial control input bounded by $u_0\in\mathcal{U}_0\subseteq\mathcal{U}$, along with Assumption 1, then, for a specified T, there exist time-varying affine abstractions $\mathcal{F}_{i,k}=\{\overline{f}_{i,k},\underline{f}_{i,k}\}$ for each system $i=\{1,2\}$ that are T-distinguishable if the following optimization problem has a solution:*

$$\min_{\substack{\overline{\Phi}_{1,k},\underline{\Phi}_{1,k},\overline{\Psi}_{1,k},\underline{\Psi}_{1,k},\\ \overline{\Omega}_{1,k},\underline{\Omega}_{1,k},\overline{\Phi}_{2,k},\underline{\Phi}_{2,k},\\ \overline{\Psi}_{2,k},\underline{\Psi}_{2,k},\overline{\Omega}_{2,k},\underline{\Omega}_{2,k},\\ \Pi,\epsilon,\alpha_1,\alpha_2}} -\epsilon+d(\alpha_1+\alpha_2) \tag{15}$$

subject to:

$$\forall((x_s,u_s)\in\mathcal{M}, x_c\in\mathcal{X}_c, u_c\in\mathcal{U}_c, i\in\{1,2\}, j\in\mathbb{Z}_1^n, k\in\mathbb{Z}_0^{T-1}):$$

$$\begin{aligned} &\underline{\Phi}_{i,k}^{(j)}x_s+\underline{\Psi}_{i,k}^{(j)}u_s+\underline{\Omega}_{i,k}^{(j)}\leq\underline{\Pi}_{1,i,k}^{(j)}f_i^{(j)}(x_s,u_s),\\ &\overline{\Phi}_{i,k}^{(j)}x_s+\overline{\Psi}_{i,k}^{(j)}u_s+\overline{\Omega}_{i,k}^{(j)}\geq\overline{\Pi}_{1,i,k}^{(j)}f_i^{(j)}(x_s,u_s),\\ &(\overline{\Phi}_{i,k}^{(j)}-\underline{\Phi}_{i,k}^{(j)})x_c+(\overline{\Psi}_{i,k}^{(j)}-\underline{\Psi}_{i,k}^{(j)})u_c+\overline{\Omega}_{i,k}^{(j)}-\underline{\Omega}_{i,k}^{(j)}\leq\alpha_i\mathbb{1},\\ &\alpha_i>0,\ \underline{\Phi}_{i,k}=\sum_{j=1}^n\underline{\Phi}_{i,k}^{(j)}, \underline{\Psi}_{i,k}=\sum_{j=1}^n\underline{\Psi}_{i,k}^{(j)}, \underline{\Omega}_{i,k}=\sum_{j=1}^n\underline{\Omega}_{i,k}^{(j)},\\ &\overline{\Phi}_{i,k}=\sum_{j=1}^n\overline{\Phi}_{i,k}^{(j)}, \overline{\Psi}_{i,k}=\sum_{j=1}^n\overline{\Psi}_{i,k}^{(j)}, \overline{\Omega}_{i,k}=\sum_{j=1}^n\overline{\Omega}_{i,k}^{(j)},\\ &\underline{\Pi}_{1,i,k}^\top=\underset{j=1}{\overset{n}{vec}}(\underline{\Pi}_{1,i,k}^{(j)}), \overline{\Pi}_{1,i,k}^\top=\underset{j=1}{\overset{n}{vec}}(\overline{\Pi}_{1,i,k}^{(j)}),\\ &\underline{\Pi}_{1,i,k}>0, \overline{\Pi}_{1,i,k}>0, \Pi_2\geq 0, \Pi_3\geq 0, \Pi_4\geq 0, \Pi_2^\top\mathbb{1}=1,\\ &\Xi+\Pi_2^\top R_2+\Pi_3^\top C=0, \Lambda+\Pi_4^\top G=0,\\ &\sum_{k=0}^{T-1}\sum_{i=1}^2(\overline{\Omega}_{i,k}-\underline{\Omega}_{i,k})\leq-\epsilon-\Pi_1^\top(\sigma+\eta_w\mathbb{1})-\Pi_2^\top r_2\\ &\qquad\qquad-\Pi_3^\top c-\Pi_4^\top g,\\ &\Pi^\top=\left[\underline{\Pi}_{1,1}^\top,\overline{\Pi}_{1,1}^\top,\underline{\Pi}_{1,2}^\top,\overline{\Pi}_{1,2}^\top,\Pi_2^\top,\Pi_3^\top,\Pi_4^\top\right], \end{aligned} \tag{16a}$$

where σ is defined in Proposition 1 and Ξ, Λ, R_2, C, G, r_2, σ, c and g are defined in the Appendix.

Proof. Analyzing the dimensions of Π in (10b), it is evident that it is a column vector. Breaking it down into the form:

$$\begin{aligned}\Pi^\top &= \left[\Pi_1^\top, \Pi_2^\top, \Pi_3^\top, \Pi_4^\top\right], \\ &= \left[\underline{\Pi}_{1,1}^\top, \overline{\Pi}_{1,1}^\top, \underline{\Pi}_{1,2}^\top, \overline{\Pi}_{1,2}^\top, \Pi_2^\top, \Pi_3^\top, \Pi_4^\top\right],\end{aligned}$$

where each $\{\underline{\Pi}_{1,i}^\top, \overline{\Pi}_{1,i}^\top\}$ for $i \in \{1,2\}$ can be further expanded as follows:

$$\begin{aligned}\underline{\Pi}_{1,i}^\top &= [\underline{\Pi}_{1,i,0}^\top, \cdots, \underline{\Pi}_{1,i,T-1}^\top], \underline{\Pi}_{1,i,k} \in \mathbb{R}^n, \ \forall k \in \mathbb{Z}_0^{T-1}, \\ \overline{\Pi}_{1,i}^\top &= [\overline{\Pi}_{1,i,0}^\top, \cdots, \overline{\Pi}_{1,i,T-1}^\top], \overline{\Pi}_{1,i,k} \in \mathbb{R}^n, \ \forall k \in \mathbb{Z}_0^{T-1}.\end{aligned} \tag{17}$$

Then, expanding the first constraint in (10b), we obtain:

$$\Pi_1^\top \bar{r}_1 + \Pi_2^\top r_2 + \Pi_3^\top c + \Pi_4^\top g \leq -\epsilon. \tag{18}$$

In (18), the first term is the only term that is bilinear. Taking it and expanding it further using matrix definitions from the Appendix, it takes the following form:

$$\Pi_1^\top \bar{r}_1 = \sum_{k=0}^{T-1} \sum_{i=1}^{2} (\overline{\Pi}_{1,i,k}^\top \overline{h}_{i,k} - \underline{\Pi}_{1,i,k}^\top \underline{h}_{i,k}) + \Pi_1^\top(\sigma + \eta_w \mathbb{1}).$$

Observing the two bilinear terms in the above equation, let:

$$\overline{\Omega}_{i,k} \triangleq \overline{\Pi}_{1,i,k}^\top \overline{h}_{i,k}, \ \underline{\Omega}_{i,k} \triangleq \underline{\Pi}_{1,i,k}^\top \underline{h}_{i,k}.$$

With the above change of variables, the algebraic manipulation of the inequality (18) results in:

$$\begin{aligned}\sum_{k=0}^{T-1} \sum_{i=1}^{2} (\overline{\Omega}_{i,k} - \underline{\Omega}_{i,k}) + \Pi_1^\top(\sigma + \eta_w \mathbb{1}) + \Pi_2^\top r_2 + \Pi_3^\top c \\ + \Pi_4^\top g \leq -\epsilon,\end{aligned} \tag{19}$$

which is linear in all the variables involved.

Similarly evaluating the second constraint in (10b) with the following change of variables:

$$\begin{aligned}\overline{\Phi}_{i,k} &\triangleq \overline{\Pi}_{1,i,k}^\top \overline{A}_{i,k}, \ \underline{\Phi}_{i,k} \triangleq \underline{\Pi}_{1,i,k}^\top \underline{A}_{i,k}, \\ \overline{\Psi}_{i,k} &\triangleq \overline{\Pi}_{1,i,k}^\top \overline{B}_{i,k}, \ \underline{\Psi}_{i,k} \triangleq \underline{\Pi}_{1,i,k}^\top \underline{B}_{i,k},\end{aligned}$$

we obtain the following set of inequalities equivalent to the second constraint in (10b):

$$\Xi + \Pi_2^\top R_2 + \Pi_3^\top C = 0, \ \Lambda + \Pi_4^\top G = 0, \ \Pi_2^\top \mathbb{1} = 1, \tag{20}$$

where Ξ and Λ are given in the Appendix.

Although originally Π_1, being a part of Π, is non-negative in the problem (10), but due of the proposed change of variables, to avoid trivial solutions, we additionally require Π_1 to be strictly positive, making our approach only sufficient.

Furthermore, based on Equations (17), let us represent each scalar component of $\underline{\Pi}_{1,i,k}$ and $\overline{\Pi}_{1,i,k}$ as $\underline{\Pi}_{1,i,k}^{(j)}$ and $\overline{\Pi}_{1,i,k}^{(j)}$ respectively, where $j \in \mathbb{Z}_1^n$. Using these scalars to scale the abstraction constraints in (10a) for each row $j \in \mathbb{Z}_1^n$ and $i \in \{1,2\}$, we have

$$\begin{aligned}\overline{\Pi}_{1,i,k}^{(j)} \overline{A}_{i,k}^{(j)} x_s + \overline{\Pi}_{1,i,k}^{(j)} \overline{B}_{i,k}^{(j)} u_s + \overline{\Pi}_{1,i,k}^{(j)} \overline{h}_{i,k}^{(j)} \\ \geq \overline{\Pi}_{1,i,k}^{(j)} f_i^{(j)}(x_s, u_s), \\ \underline{\Pi}_{1,i,k}^{(j)} \underline{A}_{i,k}^{(j)} x_s + \underline{\Pi}_{1,i,k}^{(j)} \underline{B}_{i,k}^{(j)} u_s + \underline{\Pi}_{1,i,k}^{(j)} \underline{h}_{i,k}^{(j)} \\ \leq \underline{\Pi}_{1,i,k}^{(j)} f_i^{(j)}(x_s, u_s).\end{aligned} \tag{21}$$

Replacing the bilinear terms with:

$$\begin{aligned}\overline{\Phi}_{i,k}^{(j)} &\triangleq \overline{\Pi}_{1,i,k}^{(j)} \overline{A}_{i,k}^{(j)}, \ \underline{\Phi}_{i,k}^{(j)} \triangleq \underline{\Pi}_{1,i,k}^{(j)} \underline{A}_{i,k}^{(j)}, \\ \overline{\Psi}_{i,k}^{(j)} &\triangleq \overline{\Pi}_{1,i,k}^{(j)} \overline{B}_{i,k}^{(j)}, \ \underline{\Psi}_{i,k}^{(j)} \triangleq \underline{\Pi}_{1,i,k}^{(j)} \underline{B}_{i,k}^{(j)},\end{aligned}$$

the constraints (21) take the following form:

$$\begin{aligned}\underline{\Phi}_{i,k}^{(j)} x_s + \underline{\Psi}_{i,k}^{(j)} u_s + \underline{\Omega}_{i,k}^{(j)} \leq \underline{\Pi}_{1,i,k}^{(j)} f_i^{(j)}(x_s, u_s), \\ \overline{\Phi}_{i,k}^{(j)} x_s + \overline{\Psi}_{i,k}^{(j)} u_s + \overline{\Omega}_{i,k}^{(j)} \geq \overline{\Pi}_{1,i,k}^{(j)} f_i^{(j)}(x_s, u_s).\end{aligned} \tag{22}$$

For the third constraint in (10a) that governs abstraction tightness, we start by breaking it down into two sets of sufficient constraints for $i \in \{1,2\}$:

$$\overline{A}_{i,k}^{(j)} x_c + \overline{B}_{i,k}^{(j)} u_c + \overline{h}_{i,k}^{(j)} \leq a\theta_i \mathbb{1}, \tag{23a}$$

$$-\underline{A}_{i,k}^{(j)} x_c - \underline{B}_{i,k}^{(j)} u_c - \underline{h}_{i,k}^{(j)} \leq (1-a)\theta_i \mathbb{1}, \tag{23b}$$

for some arbitrary scalar a satisfying $0 < a < 1$. Multiplying j-th row of (23a) and (23b) with $\overline{\Pi}_{1,i,k}^{(j)}$ and $\underline{\Pi}_{1,i,k}^{(j)}$, respectively, and then adding them, the following expression is obtained:

$$\begin{aligned}(\overline{\Phi}_{i,k}^{(j)} - \underline{\Phi}_{i,k}^{(j)}) x_c + (\overline{\Psi}_{i,k}^{(j)} - \underline{\Psi}_{i,k}^{(j)}) u_c + (\overline{\Omega}_{i,k}^{(j)} - \underline{\Omega}_{i,k}^{(j)}) \\ \leq (a\overline{\Pi}_{1,i,k}^{(j)} + (1-a)\underline{\Pi}_{1,i,k}^{(j)})\theta_i \mathbb{1}.\end{aligned}$$

As $a > 0$, $(1-a) > 0$, $\overline{\Pi}_{1,i,k}^{(j)} > 0$, $\underline{\Pi}_{1,i,k}^{(j)} > 0$ and θ_i, we can use $\alpha_i \triangleq (a\overline{\Pi}_{1,i,k}^{(j)} + (1-a)\underline{\Pi}_{1,i,k}^{(j)})\theta_i$ where $\alpha_i > 0$, resulting in the following form:

$$(\overline{\Phi}_{i,k}^{(j)} - \underline{\Phi}_{i,k}^{(j)}) x_c + (\overline{\Psi}_{i,k}^{(j)} - \underline{\Psi}_{i,k}^{(j)}) u_c + (\overline{\Omega}_{i,k}^{(j)} - \underline{\Omega}_{i,k}^{(j)}) \leq \alpha_i \mathbb{1}. \tag{24}$$

Lastly, $\overline{\Phi}_{i,k}$, $\underline{\Phi}_{i,k}$, $\overline{\Psi}_{i,k}$, $\underline{\Psi}_{i,k}$, $\overline{\Omega}_{i,k}$, $\underline{\Omega}_{i,k}$ also algebraically form the following relationships:

$$\begin{aligned}\overline{\Phi}_{i,k} &= \overline{\Pi}_{1,i,k}^\top \overline{A}_{i,k} = \sum_{j=1}^{n} \overline{\Pi}_{1,i,k}^{(j)} \overline{A}_{i,k}^{(j)} = \sum_{j=1}^{n} \overline{\Phi}_{i,k}^{(j)}, \\ \underline{\Phi}_{i,k} &= \underline{\Pi}_{1,i,k}^\top \underline{A}_{i,k} = \sum_{j=1}^{n} \underline{\Pi}_{1,i,k}^{(j)} \underline{A}_{i,k}^{(j)} = \sum_{j=1}^{n} \underline{\Phi}_{i,k}^{(j)}, \\ \overline{\Psi}_{i,k} &= \overline{\Pi}_{1,i,k}^\top \overline{B}_{i,k} = \sum_{j=1}^{n} \overline{\Pi}_{1,i,k}^{(j)} \overline{B}_{i,k}^{(j)} = \sum_{j=1}^{n} \overline{\Psi}_{i,k}^{(j)}, \\ \underline{\Psi}_{i,k} &= \underline{\Pi}_{1,i,k}^\top \underline{B}_{i,k} = \sum_{j=1}^{n} \underline{\Pi}_{1,i,k}^{(j)} \underline{B}_{i,k}^{(j)} = \sum_{j=1}^{n} \underline{\Psi}_{i,k}^{(j)}, \\ \overline{\Omega}_{i,k} &= \overline{\Pi}_{1,i,k}^\top \overline{h}_{i,k} = \sum_{j=1}^{n} \overline{\Pi}_{1,i,k}^{(j)} \overline{h}_{i,k}^{(j)} = \sum_{j=1}^{n} \overline{\Omega}_{i,k}^{(j)}, \\ \underline{\Omega}_{i,k} &= \underline{\Pi}_{1,i,k}^\top \underline{h}_{i,k} = \sum_{j=1}^{n} \underline{\Pi}_{1,i,k}^{(j)} \underline{h}_{i,k}^{(j)} = \sum_{j=1}^{n} \underline{\Omega}_{i,k}^{(j)}.\end{aligned} \tag{25}$$

Combining (19), (20), (22), (24) and (25), along with the relevant bounds on the decision variables, the constraints (16a) for the proposed design problem are formed.

The cost function, after replacing the terms involving θ_i with α_i, then becomes:

$$D(\epsilon, \alpha_1, \alpha_2) = -\epsilon + d(\alpha_1 + \alpha_2).$$

□

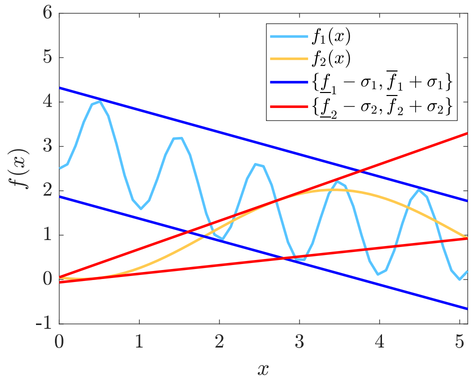


Fig. 1: Abstraction obtained for the 1D dynamics for $k = 2$.

If the solution of (15) exists, the parameters of the abstractions can be extracted using the following equations:

$$\overline{A}^{(j)}_{i,k} = \overline{\Phi}^{(j)}_{i,k}/\overline{\Pi}^{(j)}_{1,i,k}, \overline{B}^{(j)}_{i,k} = \overline{\Psi}^{(j)}_{i,k}/\overline{\Pi}^{(j)}_{1,i,k}, \overline{h}^{(j)}_{i,k} = \overline{\Omega}^{(j)}_{i,k}/\overline{\Pi}^{(j)}_{1,i,k},$$
$$\underline{A}^{(j)}_{i,k} = \underline{\Phi}^{(j)}_{i,k}/\underline{\Pi}^{(j)}_{1,i,k}, \underline{B}^{(j)}_{i,k} = \underline{\Psi}^{(j)}_{i,k}/\underline{\Pi}^{(j)}_{1,i,k}, \underline{h}^{(j)}_{i,k} = \underline{\Omega}^{(j)}_{i,k}/\underline{\Pi}^{(j)}_{1,i,k},$$

where $\overline{\Pi}^{(j)}_{1,i,k}$ and $\underline{\Pi}^{(j)}_{1,i,k}$ are scalars.

On its own, the existence of a solution to the proposed dual sufficient problem in (15) can guarantee that the obtained abstractions $\mathcal{H}_i$, $i \in \{1, 2\}$ of the two nonlinear models are distinguishable within the specified T, solving Problem 1.1. Then, by performing a line search over T, abstractions $\mathcal{H}_i$, $i \in \{1, 2\}$ that are T-distinguishable with the minimum guaranteed detection time T can be obtained, solving Problem 1.

Remark 2. *Starting from $T = 1$, incorporating a line search on T and solving the design problem in* (15) *for each increasing T, abstractions $\mathcal{H}_i$ with minimum T-detectability can be found, solving Problem 1.*

V. Simulation Examples

We apply our proposed approach on some illustrative examples, where all optimization problems are implemented in MATLAB® 2020b and solved using GUROBI 9.0.3 [27].

A. 1D Model Pair

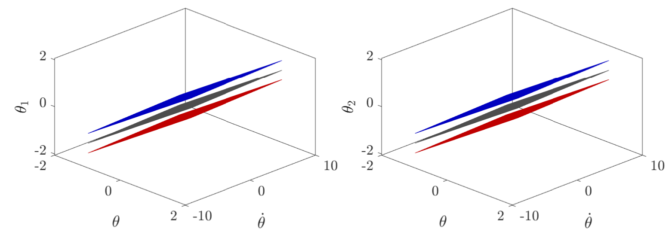
In the first example, a pair of illustrative discrete-time 1D models are considered:

$$x_{k+1} = f_1(x_k) = 1 + 0.1(x_k - 5)^2 - cos(2\pi(x - 5)),$$
$$x_{k+1} = f_2(x_k) = 1 + \frac{(x_k + 6)^2}{4000} - cos(x + 6),$$

where the state for both models is bounded by $x \in [0, 5.1]$. The process and measurement noise bounds are set as $\eta_w = 0.1$ and $\eta_v = 0.01$. Evaluating the abstractions of the given functions using Theorem 1, with a line search over T with the aim to have the smallest guaranteed detection time T, results in the model abstractions as shown in Figure 1 with the abstractions being 3-distinguishable, i.e., with $T_{min} = 3$.

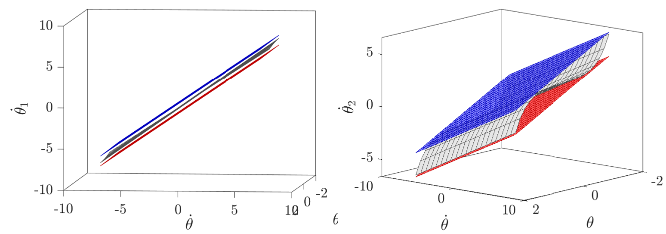
B. 2D Single-Arm Manipulator Model Pair

For the second example, the nonlinear models taken into consideration are different forms of a single-arm manipulator system with the states as $x = [\theta, \dot{\theta}]^\top$, representing the angular position and angular velocity of the arm respectively. The first model is a simple pendulum with a discrete-time model given as:



(a) Abstraction of θ_1 (b) Abstraction of θ_2

Fig. 2: Abstractions for the first state of the 2D dynamics. The (middle) white surface is the function f_1, the (top) blue plane is $\overline{\mathcal{H}}_1 + \sigma_1$ and the (bottom) red plane is $\underline{\mathcal{H}}_1 - \sigma_1$.



(a) Abstraction of $\dot{\theta}_1$ (b) Abstraction of $\dot{\theta}_2$

Fig. 3: Abstractions for the second state of the 2D dynamics. The (middle) white surface is the function f_2, the (top) blue plane is $\overline{\mathcal{H}}_2 + \sigma_2$ and the (bottom) red plane is $\underline{\mathcal{H}}_2 - \sigma_2$.

$$x_{k+1} = \begin{bmatrix} \theta_k + \dot{\theta}_k T_s \\ \dot{\theta}_k - (5.28 sin\theta_k + 7.9\dot{\theta}_k e^{-|\theta_k|} + 1.32\dot{\theta}_k)T_s \end{bmatrix},$$

where $T_s = 0.01$ is the sampling interval/period. For the second model, a 1-link robot arm with the following dynamics is considered:

$$x_{k+1} = \begin{bmatrix} \theta_k + \dot{\theta}_k T_s \\ \dot{\theta}_k + (10^{-5} cos\theta_k - 10\theta_k e^{2|\theta_k|} - 60\dot{\theta}_k)T_s \end{bmatrix},$$

where $T_s = 0.01$ is the sampling interval/period. For both models, the states are bounded by $x \in [-\pi/2, \pi/2] \times [-7, 7]$ and initial state bounds are $x \in [-\pi/4, \pi/4] \times [-3.5, 3.5]$. The process and measurement noise bounds are set at $\|w\|_\infty, \|v\|_\infty \leq 0.1$. Solving the design problem (15) with increasing values of T starting from $T = 1$, the abstractions obtained are shown in Figures 2 and 3, with the models being 1-detectable, i.e., $T_{min} = 1$.

VI. Conclusions

This paper proposed an optimization-based approach to solve the problem of finding abstractions of two functions that are guaranteed to be distinct from each other. The notion of T-distinguishability is used to enforce the distinction and the resulting algorithm is a sufficient linear programming problem, from which the abstractions can be obtained. Our proposed approach allows for linearization/model reduction of a class of systems that have different nonlinear behavioral models for easy distinction. Viable paths to explore in the future are to consider nonlinear system outputs, as opposed to linear outputs considered in this paper, as well as to consider piecewise abstractions for each model.

REFERENCES

[1] P. Tabuada, *Verification and control of hybrid systems: a symbolic approach*. Springer, 2009.

[2] A. Girard and S. Martin, "Synthesis for constrained nonlinear systems using hybridization and robust controller on symplices," *IEEE Trans. on Automatic Control*, vol. 57, no. 4, pp. 1046–1051, 2012.

[3] K. R. Singh, Q. Shen, and S. Z. Yong, "Mesh-based affine abstraction of nonlinear systems with tighter bounds," in *IEEE Conference on Decision and Control (CDC)*, 2018, pp. 3056–3061.

[4] S. Coogan and M. Arcak, "Efficient finite abstraction of mixed monotone systems," in *Proceedings of the 18th International Conference on Hybrid Systems: Computation and Control*, 2015, pp. 58–67.

[5] G. Pola, A. Girard, and P. Tabuada, "Approximately bisimilar symbolic models for nonlinear control systems," *Automatica*, vol. 44, no. 10, pp. 2508–2516, 2008.

[6] V. Alimguzhin, F. Mari, I. Melatti, I. Salvo, and E. Tronci, "Linearizing discrete-time hybrid systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5357–5364, 2017.

[7] Q. Shen and S. Z. Yong, "Robust optimization-based affine abstractions for uncertain affine dynamics," in *American Control Conference (ACC)*. IEEE, 2019, pp. 2452–2457.

[8] Z. Jin, Q. Shen, and S. Z. Yong, "Optimization-based approaches for affine abstraction and model discrimination of uncertain nonlinear systems," in *IEEE Conference on Decision and Control*, 2019.

[9] G. Reissig, "Computing abstractions of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2583–2598, 2011.

[10] T. Dang, O. Maler, and R. Testylier, "Accurate hybridization of nonlinear systems," in *ACM International Conference on Hybrid Systems: Computation and Control*, 2010, pp. 11–20.

[11] S. M. Hassaan, M. Khajenejad, S. Jensen, Q. Shen, and S. Z. Yong, "Incremental affine abstraction of nonlinear systems," *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 653–658, 2020.

[12] S. Sadraddini and C. Belta, "Formal guarantees in data-driven model identification and control synthesis," in *International Conference on Hybrid Systems: Computation and Control*. ACM, 2018, pp. 147–156.

[13] Z. Jin, M. Khajenejad, and S. Yong, "Data-driven model invalidation for unknown Lipschitz continuous systems via abstraction," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 2975–2980.

[14] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "A review of process fault detection and diagnosis: Part i: Quantitative model-based methods," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 293–311, 2003.

[15] F. Harirchi, S. Yong, and N. Ozay, "Guaranteed fault detection and isolation for switched affine models," in *IEEE Conference on Decision and Control*. IEEE, 2017, pp. 5161–5167.

[16] F. D. Bianchi and R. S. Sánchez-Peña, "Robust identification/invalidation in an LPV framework," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 3, pp. 301–312, 2010.

[17] M. Sznaier and M. C. Mazzaro, "An LMI approach to control-oriented identification and model (in)validation of LPV systems," *IEEE Trans. on Automatic Control*, vol. 48, no. 9, pp. 1619–1624, 2003.

[18] S. Prajna, "Barrier certificates for nonlinear model validation," *Automatica*, vol. 42, no. 1, pp. 117–126, 2006.

[19] N. Ozay, M. Sznaier, and C. Lagoa, "Convex certificates for model (in)validation of switched affine systems with unknown switches," *IEEE Trans. on Autom. Contr.*, vol. 59, no. 11, pp. 2921–2932, 2014.

[20] F. Harirchi, S. Z. Yong, and N. Ozay, "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*. Springer, 2018, pp. 227–251.

[21] H. Lou and R. Yang, "Conditions for distinguishability and observability of switched linear systems," *Nonlinear Analysis: Hybrid Systems*, vol. 5, pp. 427–445, 2011.

[22] E. Santis and M. Benedetto, "Observability and diagnosability of finite state systems: A unifying framework," *Automatica*, vol. 81, pp. 115–122, 2017.

[23] N. Ramdani, L. Trave-Massuyes, and C. Jauberthie, "Mode discernibility and bounded-error state estimation for nonlinear hybrid systems," *Automatica*, vol. 91, pp. 118–125, 2018.

[24] R. Niu, S. M. Hassaan, L. Yang, Z. Jin, and S. Z. Yong, "Model discrimination of switched nonlinear systems with temporal logic-constrained switching," *IEEE Control Systems Letters*, 2021.

[25] D. Bertsimas, D. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM review*, vol. 53, no. 3, pp. 464–501, 2011.

[26] M. Stämpfle, "Optimal estimates for the linear interpolation error for simplices," *Jour. of Approximation Theory*, vol. 103, pp. 78–90, 2000.

[27] Gurobi Optimization, Inc., "Gurobi optimizer reference manual," 2015. [Online]. Available: http://www.gurobi.com

APPENDIX

Time-stacking the expressions in (6b) results in the following terms for $i \in \{1, 2\}$:

$$\underline{M}_i = \begin{bmatrix} \underline{A}_{i,0} & -I_n & 0 & \cdots & 0 \\ 0 & \underline{A}_{i,1} & -I_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \underline{A}_{i,T-1} & -I_n \end{bmatrix}, p = \begin{bmatrix} p_0 \\ \vdots \\ p_T \end{bmatrix},$$

$$\underline{N}_i = \begin{bmatrix} \underline{B}_{i,0} & 0 & \cdots & 0 \\ 0 & \underline{B}_{i,1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \underline{B}_{i,T-1} \end{bmatrix}, \underline{F}_i = \begin{bmatrix} \underline{h}_{i,0} \\ \vdots \\ \underline{h}_{i,T-1} \end{bmatrix},$$

$$\overline{M}_i = \begin{bmatrix} -\overline{A}_{i,0} & I_n & 0 & \cdots & 0 \\ 0 & -\overline{A}_{i,1} & I_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\overline{A}_{i,T-1} & I_n \end{bmatrix}, q = \begin{bmatrix} q_0 \\ \vdots \\ q_{T-1} \end{bmatrix},$$

$$\overline{N}_i = \begin{bmatrix} \overline{B}_{i,0} & 0 & \cdots & 0 \\ 0 & \overline{B}_{i,1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \overline{B}_{i,T-1} \end{bmatrix}, \overline{F}_i = \begin{bmatrix} -\overline{h}_{i,0} \\ \vdots \\ -\overline{h}_{i,T-1} \end{bmatrix},$$

$$P = \begin{bmatrix} P_0 & 0 & \cdots & 0 \\ 0 & P_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & P_T \end{bmatrix}, Q = \begin{bmatrix} Q_0 & 0 & \cdots & 0 \\ 0 & Q_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & Q_{T-1} \end{bmatrix}.$$

Further, stacking of the inequalities in (11) results in the following matrices and vectors:

$$R_1 = \begin{bmatrix} \underline{M}_1 & 0 \\ \overline{M}_1 & 0 \\ 0 & \underline{M}_2 \\ 0 & \overline{M}_2 \end{bmatrix}, S_1 = \begin{bmatrix} \underline{N}_1 & 0 \\ \overline{N}_1 & 0 \\ 0 & \underline{N}_2 \\ 0 & \overline{N}_2 \end{bmatrix}, r_1 = \begin{bmatrix} -\underline{F}_1 \\ -\overline{F}_1 \\ -\underline{F}_2 \\ -\overline{F}_2 \end{bmatrix},$$

$$R_2 = \begin{bmatrix} I_{nT} & -I_{nT} \\ -I_{nT} & I_{nT} \end{bmatrix}, r_2 = 2\eta_v \mathbb{1},$$

$$C = \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix}, G = \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix}, c = \begin{bmatrix} p \\ p \end{bmatrix}, g = \begin{bmatrix} q \\ q \end{bmatrix}.$$

Finally, expanding $\Pi_1^\top R_1$ and $\Pi_1^\top S_1$ in the second constraint in (10b) yields the following row vectors, respectively, after substituting the appropriate change of variables introduced in Theorem 1:

$$\Pi_1^\top R_1 = \Xi = \begin{bmatrix} (\underline{\Xi}_1 - \overline{\Xi}_1) & (\underline{\Xi}_2 - \overline{\Xi}_2) \end{bmatrix},$$
$$\Pi_1^\top S_1 = \Lambda = \begin{bmatrix} (\underline{\Lambda}_1 - \overline{\Lambda}_1) & (\underline{\Lambda}_2 - \overline{\Lambda}_2) \end{bmatrix},$$

where, for $i \in \{1, 2\}$:

$$\underline{\Xi}_i = \begin{bmatrix} \underset{k=0}{\overset{T-1}{vec}}(\underline{\Phi}_{i,k}) & 0 \end{bmatrix} - \begin{bmatrix} 0 & \underset{k=0}{\overset{T-1}{vec}}(\underline{\Pi}_{1,i,k}^\top) \end{bmatrix}, \underline{\Lambda}_i = \underset{k=0}{\overset{T-1}{vec}}(\underline{\Psi}_{i,k}),$$
$$\overline{\Xi}_i = \begin{bmatrix} \underset{k=0}{\overset{T-1}{vec}}(\overline{\Phi}_{i,k}) & 0 \end{bmatrix} - \begin{bmatrix} 0 & \underset{k=0}{\overset{T-1}{vec}}(\overline{\Pi}_{1,i,k}^\top) \end{bmatrix}, \overline{\Lambda}_i = \underset{k=0}{\overset{T-1}{vec}}(\overline{\Psi}_{i,k}).$$