

# Combined Learning of Linear Parameter-Varying Models and Robust Control Invariant Sets

Sampath Kumar Mulagaleti and Alberto Bemporad, *Fellow, IEEE*

**Abstract**—Dynamical models identified from data are frequently employed in control system design. However, decoupling system identification from controller synthesis can result in situations where no suitable controller exists after a model has been identified. In this work, we introduce a novel control-oriented regularization in the identification procedure to ensure the existence of a controller that can enforce constraints on system variables robustly. The combined identification algorithm includes: (i) the concurrent learning of an uncertain model and a nominal model using an observer; (ii) a regularization term on the model parameters defined as the size of the largest robust control invariant set for the uncertain model. To make the learning problem tractable, we consider nonlinear models in quasi Linear Parameter-Varying (qLPV) form, utilizing a novel scheduling function parameterization that facilitates the derivation of an associated uncertain linear model. The robust control invariant set is represented as a polytope, and we adopt novel results from polytope geometry to derive the regularization function as the value of an optimization problem. Additionally, we present new model-reduction approaches that exploit the chosen model structure. Numerical examples on classical identification benchmarks demonstrate the efficacy of our approach. A simple control scheme is also derived to provide an example of data-driven control of a constrained nonlinear system.

## I. INTRODUCTION

As cyber-physical systems become increasingly prevalent, there is a growing need for algorithms that can safely and effectively control them. Model-based control design, which relies on dynamical models learned from data through system identification (SysID) techniques, has proven successful in addressing this need. These models predict future system behavior, and controllers synthesized based on them optimize the predictions to achieve desired control objectives when used in closed-loop with the system. A typical control objective is to perform reference tracking while satisfying input and output constraints. Typically, SysID and controller synthesis are performed sequentially: first, get a model from data, then design a controller for it. This approach, however, can result in controller parameterizations that are suboptimal or even infeasible for a given model parameterization [1]. For instance, a nominal controller—designed without accounting for modeling errors—may lead to constraint violations when

applied to the plant generating the data. Alternatively, a robust controller to explicitly consider model mismatch might be impossible to synthesize, i.e., there might not exist a controller of a given parameterization for the identified uncertain model. Consequently, there is a need for methodologies that integrate SysID and controller synthesis into a unified framework.

Existing unifying approaches can broadly be divided into two categories:

*i)* Identification of a model and its associated uncertainty, followed by robust controller synthesis based on the identified uncertain model. Some related contributions are [2]–[6], etc., which focus on set-membership and/or robust identification;

*ii)* Concurrent uncertain system identification and robust controller synthesis, e.g., [7]–[9], etc., in which the model, uncertainty, and controller parameterizations are fixed, and parameters are optimized to minimize a control objective while enforcing model unfalsifiability. Recent reinforcement learning-based approaches, notably, [10], can also be interpreted in this framework, in which the parameters of a linear uncertain model and a robust Model Predictive Controller (MPC) are optimized. These approaches guarantee the existence of a controller (by directly identifying it) for the uncertain model, which under reasonable assumptions is a constraint-satisfying controller for the underlying plant. However, models identified using such methods might perform poorly when used for the synthesis of control schemes other than the one they are identified for. While incorporating prediction loss into the control objective can mitigate this issue, the methods are typically designed for input-state datasets rather than input-output datasets, requiring further modifications.

Assuming there exists a method to characterize the uncertainty associated with a given model, the key step towards ensuring that one can design a model-based controller satisfying constraints involves guaranteeing the existence of a robust control invariant (RCI) set [11] for that model. RCI sets are regions of the state space in which the uncertain model can be regulated indefinitely within output constraints using feasible control inputs, and computational methods to construct them mostly rely on linear, possibly uncertain, models. Therefore, when dealing with nonlinear systems, it is reasonable to consider a linear uncertain model whose trajectories include those of the nonlinear model of the system; for this reason, Linear Parameter-Varying (LPV) model classes have been considered, and in particular quasi LPV (qLPV) models, in which the scheduling parameter is a nonlinear function of the state and input vectors. Data-driven methods for LPV system identification in the prediction error minimization paradigm

The authors are with the IMT School for Advanced Studies Lucca, Piazza San Francesco 19, 55100, Lucca, Italy (s.mulagaleti, alberto.bemporad@imtlucca.it). This work was supported by the European Research Council (ERC), Advanced Research Grant COMPACT (Grant Agreement No. 101141351).

have been developed under assumptions of availability of the scheduling function [12]. Avoiding this assumption, a large number of works have been dedicated to (implicitly) estimating the state sequence from input-output data. In [13], an autoencoder was used for defining the hidden states of a nonlinear model, such as a qLPV model (see [13, Sect. 4.2]). In [14], a joint identification procedure of the scheduling function and system matrices was presented, in which intermittent states were estimated by an encoder network [15].

### A. Contributions

We present a SysID framework to identify an uncertain model of the plant, in which we introduce a *control-oriented* regularization that guarantees the existence of a constraint-satisfying RCI set for the model being identified. We parameterize a nonlinear model in qLPV form and the associated RCI set as a polytope. We then formulate the control-oriented regularization as the value function of an optimization problem, the goal of which is to compute an RCI set with minimal conservativeness. Using every model such that the objective of the regularized SysID problem is finite, we can design a constraint-satisfying controller for the underlying plant. Our interpretation of control-oriented regularization is fundamentally different from that in [16], in which it refers to a function that minimizes deviation of closed-loop performance from a reference model.

An important novel contribution of this paper is the way we parameterize the scheduling function of the qLPV model: we use a neural network with softmax output layer, so that the outputs of the network belong to the unit simplex by construction. This implies that the model matrices of the system being identified serve directly as vertices of a linear system with multiplicative uncertainty that bounds the nonlinear system. Then, an RCI set synthesized for the uncertain linear system is also RCI for the qLPV model, at the price of some possible conservativeness. To make the identification problem computationally tractable, we parameterize the RCI set as a polytope with fixed normal directions, and enforce configuration-constraints on the offsets [17]. This enables us to express invariance conditions as being jointly linear in the parameters of the model and of the set, that we exploit to formulate an optimization problem whose value serves as our control-oriented regularization.

The paper is organized as follows. In Section II, we first state the learning problem we want to tackle. Then, we introduce the notions of uncertainty derived from model mismatch and set invariance, using which we formulate the conceptual SysID problem with control-oriented regularization. In Section III, we introduce the qLPV and RCI parameterizations, using which we formulate a computer-implementable version of the problem synthesized in Section II. In Section IV, we present the concurrent-identification problem, along with schemes to identify the system orders and RCI set template, as well as an initial point for the problem. We also discuss the limitations of our approach in this section. In Section V, we first present numerical examples to illustrate the effectiveness of the proposed qLPV parameterization with some benchmark

examples. Furthermore, we test the model reduction schemes numerically. Finally, we solve the problem developed in Section III using data collected from a nonlinear mass-spring-damper system, and use this model and the associated optimal RCI set to synthesize an output tracking control scheme.

### B. Notation

The symbols  $\mathbb{R}$  and  $\mathbb{N}$  denote the set of real and natural numbers, respectively. The set  $\mathbb{I}_a^b := \{a, \dots, b\}$  is the set of indices between  $a$  and  $b$  with  $a < b$ . Given two sets  $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$ , their Minkowski sum is defined as  $\mathcal{A} \oplus \mathcal{B} := \{a + b : a \in \mathcal{A}, b \in \mathcal{B}\}$ . If  $\mathcal{A} = \{a\}$  is singleton, with a slight abuse of notation we denote  $\{a\} \oplus \mathcal{B}$  by  $a \oplus \mathcal{B}$ . The Cartesian product  $\mathcal{A} \times \dots \times \mathcal{A}$  taken  $t$ -times is denoted by  $\mathcal{A}^t$ . Given a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , we denote the set-valued map  $f(\mathcal{A}) := \{f(a) \in \mathbb{R}^m : a \in \mathcal{A}\}$ . Given a vector  $a \in \mathbb{R}^n$ ,  $|a| \in \mathbb{R}^n$  denotes the element-wise absolute value vector. Given matrices  $M_1, \dots, M_N \in \mathbb{R}^{n \times m}$ , we define their convex hull as the set of matrices  $\text{CH}\{M_i, i \in \mathbb{I}_1^N\} := \{M \in \mathbb{R}^{n \times m} : M = \sum_{i=1}^N \lambda_i M_i, \sum_{i=1}^N \lambda_i = 1, \lambda_i \geq 0\}$ .

## II. PROBLEM DEFINITION

Assume that we have input and output data measurements available by exciting the nonlinear plant

$$\mathbf{z}^+ = \mathbf{f}(\mathbf{z}, u), \quad y = \mathbf{g}(\mathbf{z}), \quad (1)$$

where  $u \in \mathbb{R}^{n_u}$  is the input,  $y \in \mathbb{R}^{n_y}$  the output,  $\mathbf{z} \in \mathbb{R}^{n_z}$  the state vector, and  $\mathbf{z}^+$  denotes the successor state after a sample step. We assume that the functions  $\mathbf{f}$  and  $\mathbf{g}$ , as well as the state dimension  $n_z$ , are unknown. Our goal is to identify a control-oriented model of (1). Precisely, we want to solve the following problem:

*Problem 1: Given a dataset  $\mathcal{D} := \{(u_t, y_t), t \in \mathbb{I}_0^{N-1}\}$  of input-output measurements collected from (1), identify a model capable of accurately predicting the plant's behavior. Moreover, the identified model should be suitable for the synthesis of a feedback controller that regulates the plant output  $y$  within a prescribed set  $\mathbb{Y}$  under input constraints  $u \in \mathbb{U}$ .*

Problem (1) describes the entire learning-based control design pipeline, i.e., system identification and constrained controller synthesis. This problem is usually tackled in a cascaded way: first identify a model that best fits the available training/test data, then synthesize a controller. However, it may happen that, after identifying the model, synthesizing a controller that robustly satisfy output constraints under limited actuation is an impossible task, due to having separated the identification and robust control concerns. The remaining part of this section is devoted to formulate a conceptual problem of *system identification with controller synthesis guarantees*.

Before proceeding further, some clarifications regarding the structural assumptions on the underlying plant are due. Firstly, as we rely strictly on input-output data in  $\mathcal{D}$ , the identified model will inherently reside within the observable subspace of (1). Secondly, as we will see later, our approach to providing controller synthesis guarantees relies on the

feasibility of a control invariant set satisfying the constraints. Consequently, the issue of controllability is handled implicitly: If the observable subspace is fully controllable, our procedure seeks to find a model that admits a feasible control invariant set, hence solving Problem 1. Conversely, if the observable subspace contains uncontrollable modes, our procedure yields a feasible solution only if these modes are stabilizable within the constraints. Otherwise, it returns infeasibility, implying that Problem 1 does not have a feasible solution.

Let us first consider the following model

$$x^+ = f(x, u), \quad \hat{y} = g(x) \quad (2)$$

of (1), where  $x \in \mathbb{R}^{n_x}$  is the state of the model. The learning problem is usually addressed by solving the optimization problem

$$\begin{aligned} \min_{f, g, x_0} \quad & \frac{1}{N} \sum_{t=0}^{N-1} \|y_t - g(x_t)\|_2^2 \\ \text{s.t.} \quad & x_{t+1} = f(x_t, u_t), \quad t \in \mathbb{I}_0^{N-1}, \end{aligned} \quad (3)$$

where  $(u_t, y_t) \in \mathcal{D}$ , possibly under  $\ell_2$ -regularization on the coefficients parameterizing  $f$  and  $g$ . While the solution to Problem (3) might render the subsequent model-based controller synthesis feasible, this is not guaranteed. To ameliorate this, we introduce a *control-oriented regularization* into Problem (3). This regularization is based on the observation that the output of (2) may not exactly match the output of (1), such that a controller synthesized using (2) may lack *robustness* against the model mismatch, potentially leading to constraint violation. We address this using the state observer model

$$z^+ = f(z, u) + \zeta(z, u, w), \quad \hat{y} = g(z), \quad (4)$$

where  $\zeta$  is the observer function, and  $w \in \mathbb{R}^{n_w}$  is the output discrepancy defined as  $w := y - g(z)$ .

We will now characterize an uncertain model using (4) that captures the modeling error in (2) based on the following assumption.

*Assumption 1:* The behaviour of system (1) can equivalently be described by the model

$$\hat{x}^+ = f(\hat{x}, u), \quad y = g(\hat{x}) + v, \quad (5)$$

where  $v \in \mathbb{V} \subseteq \mathbb{R}^{n_y}$  is an unknown measurement noise.

We emphasize that the set  $\mathbb{V}$  is not an intrinsic property of the system (1) known a priori. Rather, it characterizes the quality of the identified model  $(f, g)$  with respect to the true system. A more accurate model results in a “smaller” set  $\mathbb{V}$ . In practice,  $\mathbb{V}$  is estimated *a posteriori* in a data-driven manner: once  $(f, g)$  are fixed by solving (3),  $\mathbb{V}$  is constructed as the set containing the output residuals  $y_t - g(\hat{x}_t)$  observed during simulation. Furthermore, while we refer to  $v$  as *measurement noise* according to standard nomenclature, we emphasize that it must be accounted for when enforcing  $y \in \mathbb{Y}$ . We use the observer in (4) to estimate the state of (5). By defining the observer error as  $e := \hat{x} - z$ , the error dynamics are given by

$$e^+ = f(\hat{x}, u) - f(z, u) - \zeta(z, u, g(\hat{x}) + v - g(z)). \quad (6)$$

*Assumption 2:* The observer function  $\zeta$  renders the error dynamics in (6) robustly stable, i.e., by denoting the input

and measurement noise sequences as  $\mathbf{u}_t := (u_0, \dots, u_{t-1})$  and  $\mathbf{v}_t := (v_0, \dots, v_{t-1})$ , respectively, there exists a robust positive invariant set  $\mathcal{E} \subseteq \mathbb{R}^{n_x}$  such that

$$e_0 \in \mathcal{E} \Rightarrow e_t \in \mathcal{E}, \quad \forall t \in \mathbb{N}, \quad (7)$$

for all input and noise sequences  $\mathbf{u}_t \in \mathbb{U}^t$  and  $\mathbf{v}_t \in \mathbb{V}^t$ .

Assumption 2 states that the error remains bounded in a set  $\mathcal{E}$  for all time irrespective of the input and noise sequences. Note that under this assumption, the input sequence is not parameterized by a controller; instead, invariance is required to hold for all admissible input sequences. While this assumption might seem strong, it is routinely employed in observer-based control design, see, e.g., [18]. For example, suppose that System (5) is described by the linear time-invariant (LTI) dynamics  $\hat{x}^+ = A\hat{x} + Bu$  and  $y = C\hat{x} + v$  for some measurement noise  $v \in \mathbb{V}$ , and the observer model in (4) is described by  $z^+ = Az + Bu + K(y - Cz)$  with observer gain  $K$ . Then, the minimal robust positive invariant set

$$\mathcal{E} = \bigoplus_{t=0}^{\infty} (A - KC)^t (-K\mathbb{V}),$$

satisfies (7), and is bounded if  $\rho(A - KC) < 1$  [19]. Similar results for Lipschitz continuous nonlinear systems were presented in [20].

Towards deriving the uncertain model, we note that since Assumptions 1 and 2 ensure the existence of sets  $\mathbb{V}$  and  $\mathcal{E}$  respectively, there always exists some disturbance set  $\mathbb{W} \subseteq \mathbb{R}^{n_y}$  that encapsulates the measurement noise and observer error by satisfying the inclusion

$$\{g(\hat{x}) + v - g(z) : \hat{x} - z \in \mathcal{E}, v \in \mathbb{V}\} \subseteq \mathbb{W}. \quad (8)$$

For the LTI example presented above, the set  $\mathbb{W} = C\mathcal{E} \oplus \mathbb{V}$  satisfies (8). In the sequel, we treat  $\mathbb{W}$  as a design parameter that must satisfy the following assumption.

*Assumption 3:* The disturbance set  $\mathbb{W} \subseteq \mathbb{R}^{n_y}$  satisfies the inclusion in (8).

For a given input sequence  $\mathbf{u}_t \in \mathbb{U}^t$  and disturbance sequence  $\mathbf{w}_t \in \mathbb{W}^t$ , let us denote by  $\hat{y}(z_0, \mathbf{u}_t, \mathbf{w}_t)$  the output of (4) at time  $t$  from initial state  $z_0 \in \mathbb{R}^{n_x}$ . Then, the reachable set of outputs of (4) at time  $t$  for all disturbance sequences and a fixed input sequence is given by

$$\hat{Y}_t(z_0, \mathbf{u}_t) := \{y \in \mathbb{R}^{n_y} : \exists \mathbf{w}_t \in \mathbb{W}^t : y = \hat{y}(z_0, \mathbf{u}_t, \mathbf{w}_t)\}$$

*Proposition 1:* Suppose Assumptions 1, 2 and 3 hold. From a given initial state  $\hat{x}_0 \in \mathbb{R}^{n_x}$ , denote the output of (5) for any given input sequence  $\mathbf{u}_t \in \mathbb{U}^t$  and noise sequence  $\mathbf{v}_t \in \mathbb{V}^t$  as  $y(\hat{x}_0, \mathbf{u}_t, \mathbf{v}_t)$ . Then, the inclusion

$$y(\hat{x}_0, \mathbf{u}_t, \mathbf{v}_t) \in \hat{Y}_t(z_0, \mathbf{u}_t), \quad \forall t \in \mathbb{N} \quad (9)$$

holds if initial states  $\hat{x}_0$  and  $z_0$  satisfy  $\hat{x}_0 - z_0 \in \mathcal{E}$ .

*Proof:* Under Assumptions 1 and 2, there exists some set  $\mathbb{W} \subseteq \mathbb{R}^{n_y}$  satisfying Assumption 3. The proof concludes if for any noise sequence  $\mathbf{v}_t \in \mathbb{V}^t$ , there exists some disturbance sequence  $\mathbf{w}_t \in \mathbb{W}^t$  such that  $y(\hat{x}_0, \mathbf{u}_t, \mathbf{v}_t) = \hat{y}(z_0, \mathbf{u}_t, \mathbf{w}_t)$  holds if  $\hat{x}_0 - z_0 \in \mathcal{E}$ . Clearly, this is satisfied by the sequence

$$\begin{aligned} \tilde{\mathbf{w}}_t = & (y(\hat{x}_0, \mathbf{u}_1, \mathbf{v}_1) - g(z_0, \mathbf{u}_1, \mathbf{v}_1)), \\ & \dots, y(\hat{x}_0, \mathbf{u}_t, \mathbf{v}_t) - g(z_0, \mathbf{u}_t, \mathbf{v}_t)). \end{aligned}$$

It remains to show that  $\tilde{\mathbf{w}}_t \in \mathbb{W}^t$ . Observe that from (7), the sequence  $\tilde{\mathbf{w}}_t$  ensures that  $\hat{x}(\hat{x}_0, \mathbf{u}_t) - z(z_0, \mathbf{u}_t, \mathbf{v}_t) \in \mathcal{E}$ , and the condition on  $\mathbb{W}$  in (8) that  $\tilde{\mathbf{w}}_t \in \mathbb{W}^t$ . ■

Hence, we can define the uncertain dynamical system

$$z^+ \in f(z, u) \oplus \zeta(z, u, \mathbb{W}), \quad y \in g(z) \oplus \mathbb{W}. \quad (10)$$

This uncertain model serves as a set-valued surrogate of the true system. Its utility is grounded in the guarantee provided by Proposition 1: provided that (10) is initialized such that its state is consistent with that of (5), i.e.,  $\hat{x}_0 - z_0 \in \mathcal{E}$ , we can simulate (10) in a set-valued manner to encompass all possible future behaviors of (5) since (9) holds. Consequently, if a controller is designed to keep the reachable output set  $\hat{Y}_t(z_0, \mathbf{u}_t)$  of (10) within the safety constraints, the true system (1) is guaranteed to satisfy them. The following result can be used to guarantee the existence of a feasible control law which ensures  $\hat{Y}_t(z_0, \mathbf{u}_t) \subseteq \mathbb{Y}$  for all  $t \in \mathbb{N}$ .

*Proposition 2:* Suppose Assumptions 1, 2 and 3 hold. Suppose further that there exists a set  $X \subseteq \mathbb{R}^{n_x}$  and a control law  $\mu(\cdot) : X \rightarrow \mathbb{U}$  that satisfy the inclusion

$$f(z, \mu(z)) \oplus \zeta(z, \mu(z), \mathbb{W}) \subseteq X, \quad \forall z \in X. \quad (11a)$$

Then, if the initial state of the observer (4) satisfies  $z_0 \in X$ , initial state of the plant in (5) satisfies  $\hat{x}_0 - z_0 \in \mathcal{E}$ , and the set  $X$  satisfies the output inclusion

$$g(z) \oplus \mathbb{W} \subseteq \mathbb{Y}, \quad \forall z \in X, \quad (11b)$$

the control input  $u_t = \mu(z_t)$ , where  $z_t$  is the state of (4) at time  $t$  from initial state  $z_0$  and any arbitrary disturbance sequence  $\mathbf{w}_t \in \mathbb{W}^t$  ensures that the plant output satisfies  $y_t \in \mathbb{Y}$  for all  $t \in \mathbb{N}$ , thus guaranteeing constraint satisfaction.

*Proof:* Under Assumptions 1, 2 and 3, inclusion (9) holds for any  $\mathbf{u}_t \in \mathbb{U}^t$  and  $\mathbf{v}_t \in \mathbb{V}^t$  if  $\hat{x}_0 - z_0 \in \mathcal{E}$  as per Proposition 1. Then, given inclusion (11a), if (4) is initialized with  $z_0 \in X$ , then we have  $z_1 \in X$  with  $u_0 = \mu(z_0)$  for all disturbances  $w_0 \in \mathbb{W}$ . From (11b) and (9), this implies that the plant satisfies  $y_1 \in \mathbb{Y}$  for all  $v_1 \in \mathbb{V}$ . Assume  $z_t \in X$ , and apply  $u_t = \mu(z_t)$ . Then we have  $z_{t+1} \in X$  for any  $w_t \in \mathbb{W}$  as per condition (11a). Consequently, by (11b) and Proposition 1, the output  $y_{t+1} \in \mathbb{Y}$ . Thus,  $y_t \in \mathbb{Y}$  for all  $t \in \mathbb{N}$ . ■

The conditions in (11) enforce  $X$  to be a robust control invariant (RCI) set for the uncertain system in (10), with invariance induced by the feedback law  $u = \mu(z)$ . Since this set  $X$  satisfies the output constraints, it guarantees the existence of a feasible controller for the plant, thus solving Problem (1). While any set  $X$  satisfying (11) guarantees safety, we aim to construct one that is minimally conservative to maximize the controller's operating range. To this end, we introduce the function  $\mathbf{r}(f, \zeta, g)$  in the sequel which satisfies

$$\mathbf{r}(f, \zeta, g) < \infty \Rightarrow \exists X \text{ satisfying (11)} \quad (12a)$$

$$\text{Small } \mathbf{r}(f, \zeta, g) < \infty \Rightarrow \exists X \text{ satisfying (11)} \quad (12b)$$

with minimal conservativeness.

The function  $\mathbf{r}(f, \zeta, g)$  is designed to ensure the existence of an RCI set  $X$ , with small values indicating that the set  $X$  is minimally conservative. While (12) characterizes qualitative

requirements on the regularization function, we quantify them in the sequel. Using this function, we modify Problem (3) as

$$\begin{aligned} \min_{f, \zeta, g, x_0} \quad & \frac{1}{N} \sum_{t=0}^{N-1} \|y_t - g(x_t)\|_2^2 + \tau \mathbf{r}(f, \zeta, g) \\ \text{s.t.} \quad & x_{t+1} = f(x_t, u_t), \quad t \in \mathbb{I}_0^{N-1}. \end{aligned} \quad (13)$$

where  $\tau \geq 0$  is the regularization parameter. We label  $\mathbf{r}(f, \zeta, g)$  as *control-oriented regularization*. As per (12), any triplet  $(f, \zeta, g)$  that admits a finite value can be used to design a controller for the underlying plant, with solutions corresponding to a small value admitting controllers with minimal conservativeness. In the sequel, we present a parameterization of the objects involved in the formulation of Problem (13), using which we derive a tractable formulation of the optimization problem.

*Connections to existing literature:* Our contribution aims to identify open-loop plant models for set-based robust control. Traditional approaches for jointly characterizing a model and its associated uncertainty, such as [3], [6], [21], typically operate under conditions similar to Assumption 1 with a known set  $\mathbb{V}$ . While these methods yield a direct characterization of the uncertain (potentially nonlinear) model, they do not guarantee the existence of an RCI set, as the output constraint set  $\mathbb{Y}$  is not incorporated into the identification procedure. In learning-based control frameworks like [5], the set  $\mathbb{V}$  is estimated along with an uncertain linear model for use in robust control synthesis. However, because the constraints  $\mathbb{Y}$  are decoupled from the identification phase, the resulting control synthesis problem may still be infeasible. Recent contributions, such as [9], [22], address this by identifying a model and a set  $\mathbb{V}$  while explicitly enforcing that the model admits an RCI set within  $\mathbb{Y}$ , thus coupling system identification with RCI set synthesis. In these works, a parameterization of the RCI set is selected a priori, with [9] deriving rigorous bounds to satisfy Assumption 1. Nevertheless, these approaches are currently limited to identifying uncertain linear systems from datasets containing full state measurements. Our approach extends these ideas to the identification of uncertain nonlinear systems using only input-output data, while guaranteeing the existence of a feasible RCI set. A related line of research involves safe Reinforcement Learning (RL) with robust MPC parameterizations [10]. These methods also tackle the estimation of  $\mathbb{V}$  online, and parameterize  $X$  as the feasible region of the MPC controller. Finally, behavioral system [23] approaches, such as [24], have gained attention for synthesizing robust controllers directly from noisy input-output trajectories. While originally designed for linear systems, recent extensions aim to generalize these capabilities to nonlinear models [25], [26]. A detailed comparison of our approach with these data-driven methods is a subject for future work.

### III. MODEL AND SET PARAMETERIZATION

In this section, we present a model and RCI set parameterization, using which we formulate Problem (13). We parameterize the model as a qLPV system. Such models describe state evolution using a time-varying linear map, with the linear maps scheduled using an explicitly characterized function [27]. They

have frequently been observed to provide a very good balance between prediction accuracy and ease of robust controller design. In this work, we exploit a particular choice of scheduling function parameterization that enables a straightforward derivation of a linear system with multiplicative uncertainty that bounds the qLPV system. Then, we parameterize the RCI set as a polytope with fixed normal vectors, in which we induce robustness for the uncertain linear system using linear inequalities. Finally, we present a formulation of the control-oriented regularization function  $\mathbf{r}$  for our parameterization.

### A. Dynamical system

We parameterize model (2) as the qLPV system

$$x^+ = A(p(x, u))x + B(p(x, u))u, \quad \hat{y} = Cx, \quad (14)$$

and the state observer in (4) with  $\zeta(z, u, w) = K(p(z, u))w$ , where the matrix-valued functions  $(A(p), B(p), K(p))$  depend linearly on the scheduling vector  $p \in \mathbb{R}^{n_p}$  as

$$(A(p), B(p), K(p)) := \sum_{i=1}^{n_p} p_i (A_i, B_i, K_i). \quad (15)$$

Further, we model each component of the scheduling function  $p : \mathbb{R}^{n_x+n_u} \rightarrow \mathbb{R}^{n_p}$  as

$$p_i(x, u) = \begin{cases} \frac{e^{\mathcal{N}_i(x, u)}}{1 + \sum_{j=1}^{n_p-1} e^{\mathcal{N}_j(x, u)}}, & i \in \mathbb{I}_1^{n_p-1}, \\ \frac{1}{1 + \sum_{j=1}^{n_p-1} e^{\mathcal{N}_j(x, u)}}, & i = n_p. \end{cases} \quad (16)$$

where each  $\mathcal{N}_i : \mathbb{R}^{n_x+n_u} \rightarrow \mathbb{R}$  is a feedforward neural network (FNN) with  $n_h$  hidden layers. The parameterization (16) forces  $p$  to belong to the simplex  $\mathcal{P}$  defined as

$$\mathcal{P} := \left\{ p : \sum_{i=1}^{n_p} p_i = 1, 0 \leq p \leq 1 \right\}. \quad (17)$$

Note that (16) can be interpreted as the output of a classifier for predicting a multi-category target of dimension  $n_p$ , where  $p_i$  is the probability of the target being in category  $i$ , i.e.,  $p(x, u)$  is a feedforward neural network with softmax output. Consequently, the parameterization in (16) makes (14) a Mixture-of-Experts (MoE) model. MoE architectures are known universal approximators for continuous functions on compact sets [28]. While the structure of (14) assumes the origin to be an equilibrium point, affine dynamics can be readily accommodated by augmenting the state vector as  $\tilde{x} = [x^\top \ 1]^\top$ . Therefore, provided that the plant dynamics in (1) are continuous, the proposed qLPV parameterization is sufficiently expressive to satisfy Assumption 1.

*Remark 1:* It is important to distinguish the proposed identification approach from the ‘‘local approach’’ to LPV identification, in which independent LTI models are estimated at fixed operating points in arbitrary statespaces, requiring coherence transformations to unify them before interpolation [29]. In contrast, we employ a global identification approach, e.g., [14], using a self-scheduled parameter  $p(x, u)$  as in (16), and identify the global coefficient matrices  $(A_i, B_i, K_i)$  along with the scheduling function directly. This construction ensures that

the state  $x$  is defined in a unified basis, rendering coherence transformations unnecessary. Furthermore, identification is performed directly using an input-output dataset without requiring a scheduling variable dataset.

### B. Robust control invariant set

We parameterize the RCI set  $X \subseteq \mathbb{R}^{n_x}$  as the polytope

$$X \leftarrow \mathbb{X}(q) := \{x \in \mathbb{R}^{n_x} : Fx \leq q\},$$

where  $F \in \mathbb{R}^{f \times n_x}$  is a matrix that we fix a priori, and  $q \in \mathbb{R}^f$  is the variable that characterizes the set. Over  $q$ , we enforce *configuration-constraints* [17], which are conic constraints of the form  $\mathcal{C} := \{q : Eq \leq 0\}$ . These constraints dictate that

$$q \in \mathcal{C} \Rightarrow \mathbb{X}(q) = \text{CH}\{V_l q, l \in \mathbb{I}_1^f\}, \quad (18)$$

where  $\{V_l \in \mathbb{R}^{f \times n_x}, l \in \mathbb{I}_1^f\}$  are the vertex maps given a priori. Then, we parameterize the disturbance set as

$$\mathbb{W} \leftarrow \mathbb{W}_\kappa(c_w, \epsilon_w) := \{w \in \mathbb{R}^{n_y} : |w - c_w| \leq \kappa \epsilon_w\}, \quad (19)$$

where  $c_w, \epsilon_w \in \mathbb{R}^{n_y}$  are the parameters, and  $\kappa > 1$  is a user-specified parameter to account for finite data. We present a characterization of the disturbance set parameters in the sequel. Firstly, we enforce the set  $\mathbb{X}(q)$  to be RCI for the uncertain system in (10), written for the qLPV parameterization as

$$\begin{aligned} z^+ &\in A(p(z, u))z + B(p(z, u))u \oplus K(p(z, u))\mathbb{W}_\kappa(c_w, \epsilon_w), \\ y &\in Cz \oplus \mathbb{W}_\kappa(c_w, \epsilon_w) \end{aligned} \quad (20)$$

The following result is key to our developments.

*Proposition 3:* Let the convex hull of the matrices  $(A_i, B_i, K_i)$  be denoted by  $\Delta := \text{CH}\{(A_i, B_i, K_i), i \in \mathbb{I}_1^{n_p}\}$ . An RCI set  $\mathbb{X}(q)$  for the uncertain linear system

$$\begin{aligned} z^+ &\in Az + Bu \oplus K\mathbb{W}_\kappa(c_w, \epsilon_w), \\ y &\in Cx \oplus \mathbb{W}_\kappa(c_w, \epsilon_w) \end{aligned} \quad (21)$$

with multiplicative uncertainty  $(A, B, K) \in \Delta$  is RCI also for the uncertain nonlinear system in (20).

*Proof:* Take any  $z \in \mathbb{X}(q)$ . Since  $\mathbb{X}(q)$  is RCI for the linear system in (21), there exists some input  $u \in \mathbb{U}$  such that  $Az + Bu + Kw \in \mathbb{X}(q)$  for all  $w \in \mathbb{W}_\kappa(c_w, \epsilon_w)$  and  $(A, B, K) \in \Delta$ . From the parameterizations in (15), (16) and (17), we have that  $(A(p(z, u)), B(p(z, u)), K(p(z, u))) \in \Delta$ , which implies the subsequent state of (20) also satisfies

$$A(p(z, u))z + B(p(z, u))u + K(p(z, u))w \in \mathbb{X}(q)$$

for all  $w \in \mathbb{W}_\kappa(c_w, \epsilon_w)$ . Hence,  $\mathbb{X}(q)$  is RCI also for (20). ■

Exploiting Proposition 3, we now characterize conditions on  $q$  such that  $\mathbb{X}(q)$  is an RCI set for (21). In this result, we assume that the constraint sets  $\mathbb{U}$  and  $\mathbb{Y}$  are defined as

$$\mathbb{U} := \{u \in \mathbb{R}^{n_u} : H^u u \leq h^u\}, \quad \mathbb{Y} := \{y \in \mathbb{R}^{n_y} : H^y y \leq h^y\}.$$

*Proposition 4:* A set  $\mathbb{X}(q)$  is an RCI set for system (21) with respect to constraints  $y \in \mathbb{Y}$  and  $u \in \mathbb{U}$  if there exist control inputs  $u := (u_1, \dots, u_w)$  such that the inequalities

$$F(A_i V_l q + B_i u_l + K_i c_w) + \kappa |F K_i| \epsilon_w \leq q, \quad (22a)$$

$$H^y (C V_l q + c_w) + \kappa |H^y| \epsilon_w \leq h^y, \quad (22b)$$

$$H^u u_l \leq h^u, \quad Eq \leq 0, \quad (22c)$$

are verified for all  $i \in \mathbb{I}_1^{n_p}$  and  $l \in \mathbb{I}_1^y$ .

*Proof:* The proof follows from definition of RCI sets and [17, Corollary 4] which exploits (18), where  $u$  denotes the vertex control inputs. ■

For fixed system matrices  $(A_i, B_i, K_i, C)$  and disturbance set parameters  $(c_w, \epsilon_w, \kappa)$ , the conditions in (22) are linear in  $(q, u)$ . We exploit this property in the development of the regularization function  $r$  in the sequel.

### C. Disturbance set characterization

We recall from Proposition 2 that the RCI set  $\mathbb{X}(q)$  of (20) can be used to guarantee the existence of a controller for the underlying plant if the disturbance set  $\mathbb{W}_\kappa(c_w, \epsilon_w)$  satisfies Assumption 3. We now present a procedure to estimate the disturbance set parameters  $(c_w, \epsilon_w, \kappa)$  such that this assumption is satisfied. Assuming access to an additional dataset  $\mathcal{D}^w := \{(y_t, u_t), t \in \mathbb{I}_0^{N_w-1}\}$  of input-output measurements from the plant, we simulate the closed-loop observer

$$z^+ = A(p(z, u))z + B(p(z, u))u + K(p(z, u))(y - Cz), \quad (23)$$

(equivalent to (4)) using inputs and outputs from  $\mathcal{D}^w$ , from  $z_0 = 0$ . Then, building a sampled set of the output discrepancy as  $\mathcal{W} := \{y_t - Cz_t, t \in \mathbb{I}_0^{N_w-1}\}$ , we can define the prediction error bounds taken elementwise as  $\bar{w} := \max_{w \in \mathcal{W}} w$  and  $\underline{w} := \min_{w \in \mathcal{W}} w$ . Given the disturbance set parameterization in (19), the inclusion  $\mathcal{W} \subset \mathbb{W}_\kappa(c_w, \epsilon_w)$  then holds with

$$c_w = 0.5(\bar{w} + \underline{w}), \quad \epsilon_w = 0.5(\bar{w} - \underline{w}) \quad (24)$$

for any  $\kappa \geq 1$ . Since  $\mathcal{D}^w$  is a finite dataset however, the disturbance set  $\mathbb{W}_\kappa(c_w, \epsilon_w)$  might not satisfy Assumption 3 with  $\kappa = 1$ . In the following result, we present a lower-bound on the inflation parameter  $\kappa$  to ensure its satisfaction.

*Proposition 5:* Suppose Assumptions 1 and 2 are satisfied by the model structure in (14) and observer function  $K(p(z, u))w$  respectively, such that there exists an invariant set  $\mathcal{E} \subseteq \mathbb{R}^{n_x}$  in which the observer error  $e = \hat{x} - z$  is bounded. Suppose that  $0 \in \mathcal{E}$ ,  $\hat{x}_0 \in \mathcal{E}$ , and the dataset  $\mathcal{D}^w$  is such that

$$\forall \delta \in C\mathcal{E} \oplus \mathbb{V}, \exists t \in \mathbb{I}_0^{N_w-1} : |\delta - (y_t - Cz_t)| \leq \alpha, \quad (25)$$

for some  $\alpha > 0$ , where  $z_t$  is the state of (23) simulated from  $z_0 = 0$  using  $\mathcal{D}^w$ . Then, if

$$\kappa \geq 1 + \max_{i \in \mathbb{I}_1^{n_y}} \alpha_i / \epsilon_{w_i}, \quad (26)$$

the disturbance set  $\mathbb{W}_\kappa(c_w, \epsilon_w)$  with  $(c_w, \epsilon_w)$  estimated as in (24) satisfies Assumption 3.

*Proof:* Assumptions 1 and 2, along with  $\hat{x}_0 \in \mathcal{E}$  and  $z_0 = 0$  guarantee that  $e_t = \hat{x}_t - z_t \in \mathcal{E}$  for all  $t \in \mathbb{I}_0^{N_w-1}$ . For any plant and observer states  $\hat{x}, z \in \mathbb{R}^{n_x}$  satisfying the inclusion  $C(\hat{x} - z) \in C\mathcal{E} \oplus \mathbb{V}$ , we can then write

$$\begin{aligned} & |C(\hat{x} - z) - c_w| \\ & \leq |(y_t - Cz_t) - c_w| + |C(\hat{x} - z) - (y_t - Cz_t)| \\ & \leq \epsilon_w + \alpha \leq \kappa \epsilon_w, \end{aligned}$$

where the first inequality follows from the triangle inequality, the second from (24) and (25) for some  $t \in \mathbb{I}_0^{N_w-1}$ , and the

third from (26). Hence,  $C\mathcal{E} \oplus \mathbb{V} \subseteq \mathbb{W}_\kappa(c_w, \epsilon_w)$  holds, thus satisfying Assumption 3. ■

*Assumption 4:* The inflation parameter  $\kappa > 1$  is chosen to be large enough that (26) is verified.

Under Assumption 4, the disturbance set  $\mathbb{W}_\kappa(c_w, \epsilon_w)$  is guaranteed to satisfy Assumption 3, such that ensuring the existence of an RCI set  $\mathbb{X}(q)$  ensures the existence of a controller for the underlying plant as per Proposition 2. In practice, a priori satisfaction of Assumption 4 is difficult to guarantee as  $\alpha$  is generally unknown. While  $\kappa \approx 1$  may suffice for a large dataset  $\mathcal{D}^w$  generated by persistently exciting inputs, the exact computation of  $\kappa$  remains a fundamental challenge in data-driven control [9], [10]. We remark that  $\kappa$  acts as a *robustness parameter*: larger values yield a larger disturbance set ensuring the RCI set is robust against significant model uncertainty, but potentially leading to infeasibility. While bootstrapping techniques utilizing multiple datasets  $\mathcal{D}^w$  could offer a way to estimate  $\kappa$  more accurately, the development of such methods remains a subject for future research.

### D. Control-oriented regularization function

We now present a formulation of the regularization function  $r$  satisfying the requirements in (12). We denote by  $\Theta$  the vector containing the parameters of the qLPV models in (14) and (20), i.e., the matrices  $(A_i, B_i, K_i)$ , the output matrix  $C$ , and the weights of the FNNs parameterizing the scheduling function in (16). Given  $\Theta$ , the disturbance set parameters  $(c_w, \epsilon_w, \kappa)$  are uniquely defined as per (24). We recall that  $\Theta$  parameterizes the linear inequalities in (22). Accordingly, we define the set of admissible RCI parameters as

$$\mathbb{Q}(\Theta) := \{q \in \mathbb{R}^f : \exists u \in \mathbb{R}^{v_u} : (q, u) \text{ satisfy (22)}\}. \quad (27)$$

As established in Propositions 3 and 4, any vector  $q \in \mathbb{Q}(\Theta)$  defines an RCI set  $\mathbb{X}(q)$  for the uncertain nonlinear system (20). Hence, by Proposition 2, if  $\mathbb{Q}(\Theta)$  is nonempty under the aforementioned assumptions, the identified model can be used to design a controller for the underlying plant.

In designing the regularization function, we emphasize that the metric for the *conservativeness* of an RCI set must align with the downstream control objectives. For instance, if the goal is robust stabilization to the origin, the function

$$r(\Theta) := \min_{q \in \mathbb{Q}(\Theta)} \|q\|_1 \quad (28)$$

inspired by [30] is appropriate, as it encourages a minimal RCI set. In this paper, our objective is to synthesize output tracking controllers. To this aim, we assume access to an output reference trajectory  $\{y_k^r \in \mathbb{R}^{n_y}, k \in \mathbb{I}_0^M\}$ , using which we define the regularization function as

$$\begin{aligned} r(\Theta) & := \min_{q, v} \sum_{k=0}^M \|y_k^r - Cz_k\|_2^2 \\ & \text{s.t. } \begin{cases} q \in \mathbb{Q}(\Theta), z_0 = 0, v_k \in \mathbb{U}, z_k \in \mathbb{X}(q), \\ z_{k+1} = A(p(z_k, v_k))z_k + B(p(z_k, v_k))v_k, \\ k \in \mathbb{I}_0^{M-1}, \end{cases} \end{aligned} \quad (29)$$

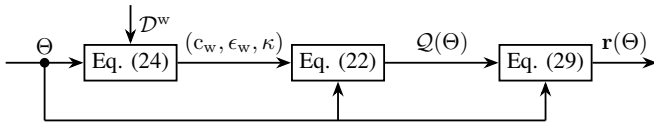


Fig. 1: Evaluation of  $r(\Theta)$  given  $\Theta$  and  $\mathcal{D}^w$ .

and define  $r(\Theta) = \infty$  if Problem (29) is infeasible. In (29), we compute an RCI set  $\mathbb{X}(q)$  and a feasible input sequence  $\mathbf{v} = (v_0, \dots, v_{M-1})$  that makes the qLPV system track the reference while staying inside the RCI set. In Figure 1, we present a schematic of the evaluation of  $r(\Theta)$ . This formulation satisfies the requirements in (12): if  $\mathbb{Q}(\Theta)$  is empty, the value is  $\infty$ ; if nonempty, a small value indicates that the open-loop system is capable of tight tracking, implying reduced conservativeness relevant to the control task. Problem (29) can be adapted for multiple reference trajectories by considering a different open-loop sequence for each reference. Due to the qLPV dynamics, it is a nonlinear optimization problem. A convex alternative can be derived by replacing the qLPV dynamics with the uncertain LTI system in (21) and employing robust tube-based MPC techniques (e.g., [31]) to propagate the set-valued dynamics under additive disturbances. The development of such relaxations is left for future research.

#### IV. REGULARIZED SYSTEM IDENTIFICATION ALGORITHM

Based on the aforementioned developments, we formulate the regularized system identification problem in (13) as

$$\begin{aligned} \min_{\Theta, x_0} \frac{1}{N} \sum_{t=0}^{N-1} \|y_t - Cx_t\|_2^2 + \tau r(\Theta) \quad (30) \\ \text{s.t. } x_{t+1} = A(p(x_t, u_t))x_t + B(p(x_t, u_t))u_t, \quad t \in \mathbb{I}_0^{N-1}, \end{aligned}$$

where  $\Theta$  denotes the model parameters,  $r(\Theta)$  is defined in (29), and  $\tau \geq 0$  is the regularization parameter. We recall that any value of  $\Theta$  resulting in  $\mathbb{Q}(\Theta)$  from (27) being nonempty will lead to  $r(\Theta) < \infty$ , such that the corresponding model can be used for constrained control synthesis for the plant. Since (29) is itself a minimization problem, we denote its constraints as  $c(\Theta, q, \mathbf{v}) \leq 0$ , and append  $(q, \mathbf{v})$  as optimization variables to Problem (30) resulting in the formulation

$$\begin{aligned} \min_{\Theta, x_0, q, \mathbf{v}} \frac{1}{N} \sum_{t=0}^{N-1} \|y_t - Cx_t\|_2^2 + \tau \sum_{k=0}^M \|y_k^r - Cz_k\|_2^2 \quad (31) \\ \text{s.t. } \begin{cases} x_{t+1} = A(p(x_t, u_t))x_t + B(p(x_t, u_t))u_t, \\ z_{k+1} = A(p(z_k, v_k))z_k + B(p(z_k, v_k))v_k, \\ z_0 = 0, \quad c(\Theta, q, \mathbf{v}) \leq 0, \quad t \in \mathbb{I}_0^{N-1}, \quad k \in \mathbb{I}_0^{M-1}. \end{cases} \end{aligned}$$

While Problem (30) can be solved using any nonlinear programming solver [32], the approach may become intractable if the datasets  $\mathcal{D}$  or  $\mathcal{D}^w$  are large. Recall that  $\mathcal{D}^w$  is used to formulate the inequalities through (24) and (22) (See Figure 1). Hence, we propose to instead penalize the inequality

constraints using a soft penalty, resulting in

$$\begin{aligned} \min_{\Theta, x_0, q, \mathbf{v}} \frac{1}{N} \sum_{t=0}^{N-1} \|y_t - Cx_t\|_2^2 + \tau \sum_{k=0}^M \|y_k^r - Cz_k\|_2^2 \quad (32) \\ + \tau_c \|\max\{c(\Theta, q, \mathbf{v}), 0\}\|_1, \end{aligned}$$

where  $\tau_c > 0$  is the penalty parameter that must be chosen high enough to ensure satisfaction of the constraints. Note that the state trajectories  $(x_0, \dots, x_N)$  and  $(z_0, \dots, z_M)$  are implicit functions of  $(x_0, \Theta)$  and  $(\mathbf{v}, \Theta)$  respectively, such that the equality constraints can be eliminated by condensing the dynamics. Since Problem (32) is unconstrained, it can be tackled with solvers such as Adam [33], L-BFGS [34], etc. While (32) always admits a finite solution, large values indicate infeasibility of (31). Note that in general, Problem (32) does not admit a unique solution, since it is a nonconvex optimization problem. In the sequel, we present approaches to identify the scheduling variable order  $n_p$ , and matrix  $F$  required to formulate Problem (32). For an approach to identify the model order  $n_x$ , we refer the reader to [35, Sec. II].

#### A. Scheduling variable order selection

We present an approach to decide the scheduling variable order  $n_p$  by solving an identification problem. Towards formulating this problem, we assume that the FNNs parameterizing the scheduling function in (16) are given by

$$\mathcal{N}_i(x, u) := W_i^L \hat{\mathcal{N}}_i(x, u) + b_i^L, \quad \forall i \in \mathbb{I}_1^{n_p-1}. \quad (33)$$

where  $W_i^L \in \mathbb{R}^{1 \times n_L}$  and  $b_i^L \in \mathbb{R}^1$  are weights and biases of the final layer of the FNNs, and  $\hat{\mathcal{N}}_i: \mathbb{R}^{n_x+n_u} \rightarrow \mathbb{R}^{n_L}$  is the part of the FNN till the final linear layer. We note that the parameter vector  $\Theta$  includes  $\{(W_i^L, b_i^L), i \in \mathbb{I}_1^{n_p-1}\}$ . Then, using the dataset  $\mathcal{D}$ , we formulate the problem

$$\begin{aligned} \min_{\Theta, x_0} \frac{1}{N} \sum_{t=0}^{N-1} \|y_t - Cx_t\|_2^2 + \kappa_p \sum_{i=1}^{n_p-1} \|W_i^{L\top}\|_2 \quad (34) \\ \text{s.t. } x_{t+1} = A(p(x_t, u_t))x_t + B(p(x_t, u_t))u_t, \quad t \in \mathbb{I}_0^{N-1}, \end{aligned}$$

where  $\kappa_p > 0$  is the regularization parameter. The regularization in (34) acts as a group-Lasso penalty, encouraging group sparsity in the vectors  $W_i^L$ . If  $W_i^L$  is driven to zero, the corresponding FNN reduces to a constant bias  $\mathcal{N}_i(x, u) = b_i^L$ . In the following result, we derive the reduced-order model under the assumption that the solution to (34) yields sparse weights. Without loss of generality, we assume the indices are ordered such that the non-zero weights appear first.

**Proposition 6:** Let the solution of (34) be such that we have  $W_i^L \neq 0$  for all  $i \in \mathbb{I}_1^{\tilde{n}_p-1}$ , and  $W_i^L = 0$  for all  $i \in \mathbb{I}_{\tilde{n}_p}^{n_p-1}$ . Let  $\beta := 1 + \sum_{i=\tilde{n}_p}^{n_p-1} e^{b_i^L}$  and  $\tilde{\mathcal{N}}_i(x, u) := \mathcal{N}_i(x, u) - \log(\beta)$ , for all  $i \in \mathbb{I}_1^{\tilde{n}_p-1}$ . Then the system in (14) identified by (34) can be equivalently expressed by the matrix functions

$$(A(p(x, u)), B(p(x, u))) = \sum_{i=1}^{\tilde{n}_p} \tilde{p}_i(x, u) (\tilde{A}_i, \tilde{B}_i), \quad (35)$$

where  $(\tilde{A}_i, \tilde{B}_i) = (A_i, B_i)$  for  $i \in \mathbb{I}_1^{\tilde{n}_p-1}$  and

$$(\tilde{A}_{\tilde{n}_p}, \tilde{B}_{\tilde{n}_p}) = \frac{1}{\beta} \left( \sum_{i=\tilde{n}_p}^{n_p-1} e^{b_i^L} (A_i, B_i) + (A_{n_p}, B_{n_p}) \right), \quad (36)$$

and the scheduling function  $\tilde{p} : \mathbb{R}^{n_x+n_u} \rightarrow \mathbb{R}^{\tilde{n}_p}$  with

$$\tilde{p}_i(x, u) = \begin{cases} \frac{e^{\tilde{\mathcal{N}}_i(x,u)}}{1 + \sum_{j=1}^{\tilde{n}_p-1} e^{\tilde{\mathcal{N}}_j(x,u)}}, & i \in \mathbb{I}_1^{\tilde{n}_p-1}, \\ 1, & i = \tilde{n}_p. \end{cases} \quad (37)$$

*Proof:* By definition, for all  $i \in \mathbb{I}_1^{\tilde{n}_p-1}$ , we have

$$p_i(x, u) = \frac{e^{\mathcal{N}_i(x,u)}}{\beta + \sum_{j=1}^{\tilde{n}_p-1} e^{\mathcal{N}_j(x,u)}} = \frac{e^{\tilde{\mathcal{N}}_i(x,u)}}{1 + \sum_{j=1}^{\tilde{n}_p-1} e^{\tilde{\mathcal{N}}_j(x,u)}},$$

which corresponds to  $\tilde{p}_i(x, u)$ . Next, substituting  $W_i^L = 0$  for  $i \in \mathbb{I}_{\tilde{n}_p}^{n_p-1}$  into the matrix function  $A(p(x, u))$ , we obtain

$$\begin{aligned} A(p(x, u)) &= \sum_{i=1}^{\tilde{n}_p-1} A_i p_i(x, u) + \frac{\sum_{i=\tilde{n}_p}^{n_p-1} e^{b_i^L} A_i + A_{n_p}}{\beta + \sum_{j=1}^{\tilde{n}_p-1} e^{\mathcal{N}_j(x,u)}} \\ &= \sum_{i=1}^{\tilde{n}_p-1} A_i \tilde{p}_i(x, u) + \frac{\beta \tilde{A}_{\tilde{n}_p}}{\beta + \sum_{j=1}^{\tilde{n}_p-1} e^{\mathcal{N}_j(x,u)}}, \end{aligned}$$

where the term multiplying  $\tilde{A}_{\tilde{n}_p}$  is  $\tilde{p}_{\tilde{n}_p}(x, u)$ . The derivation for  $B(p(x, u))$  follows identically, concluding the proof. ■

The regularization in (34) along with Proposition 6 present an approach to identify an equivalent qLPV model with scheduling variable order  $\tilde{n}_p \leq n_p$ . In some cases however, it might be desirable to reduce the order further. While this problem has been extensively studied in the literature (c.f. [36], [37] for a comprehensive review), the subsequent reduced-order functions might lose their simplex structure. Hence, we will now develop a procedure to reduce the order further by solving simple convex quadratic programs (QPs).

**A posteriori order reduction:** Given a model with scheduling variable order  $n_p$ , the procedure in Proposition 6 relies on the fact that  $W_i^L = 0$  forces the FNN output to a constant value  $\mathcal{N}_i(x, u) = b_i^L$ . This sparsity allows specific system matrices to be lumped together without loss of accuracy. We extend this intuition to an *a posteriori* reduction procedure. By examining the scheduling trajectories  $p(x_t, u_t)$  generated during simulation with the dataset  $\mathcal{D}$ , we can identify components that exhibit negligible variation. If  $p_i(x_t, u_t)$  remains nearly constant for all  $t \in \mathbb{I}_0^{N-1}$ , the corresponding scheduling variable can be eliminated to derive an approximate reduced-order qLPV model. To formalize this intuition, we assume that the scheduling variables to be eliminated correspond to indices  $i \in \mathbb{I}_{\tilde{n}_p}^{n_p-1}$  with  $\tilde{n}_p < n_p$ , and denote the reduced-order qLPV model that approximates (14) as

$$\tilde{x}^+ = \tilde{A}(\tilde{p}(\tilde{x}, u))\tilde{x} + \tilde{B}(\tilde{p}(\tilde{x}, u))u, \quad \tilde{y} = \tilde{C}\tilde{x}, \quad (38)$$

where the matrix-valued functions are parameterized as

$$(\tilde{A}(\tilde{p}), \tilde{B}(\tilde{p})) := \sum_{i=1}^{\tilde{n}_p} \tilde{p}_i(\tilde{A}_i, \tilde{B}_i), \quad (39)$$

and each component of the reduced-order scheduling function  $\tilde{p} : \mathbb{R}^{n_x+n_u} \rightarrow \mathbb{R}^{\tilde{n}_p}$  is parameterized as

$$\tilde{p}_i(\tilde{x}, u) := \begin{cases} \frac{e^{W_i^L \hat{\mathcal{N}}_i(\tilde{x}, u)} e^{\tilde{b}_i^L}}{1 + \sum_{j=1}^{\tilde{n}_p-1} e^{W_j^L \hat{\mathcal{N}}_j(\tilde{x}, u)} e^{\tilde{b}_j^L}}, & i \in \mathbb{I}_1^{\tilde{n}_p-1}, \\ 1, & i = \tilde{n}_p. \end{cases} \quad (40)$$

The parameters of (38) are  $\{\tilde{A}_i, \tilde{B}_i, i \in \mathbb{I}_1^{\tilde{n}_p}\}$ ,  $\tilde{C}$ , and biases  $\{\tilde{b}_i^L, i \in \mathbb{I}_1^{\tilde{n}_p-1}\}$  of the last layers of the FNNs, with the weights of the previous layers retained the same as (14). The reduction procedure computes these parameters to achieve

$$(A(p(x_t, u_t)), B(p(x_t, u_t))) \approx (\tilde{A}(\tilde{p}(\tilde{x}_t, u_t)), \tilde{B}(\tilde{p}(\tilde{x}_t, u_t))), \quad (41)$$

where  $x_t$  and  $\tilde{x}_t$  are the state trajectories of (14) and (38) when simulated from the same initial state, e.g., obtained by solving Problem (34), using inputs from the dataset  $\mathcal{D}$ . To this end, we make the following assumption.

**Assumption 5:** The components  $e^{W_i^L \hat{\mathcal{N}}_i(x_t, u_t)}$  for  $i \in \mathbb{I}_{\tilde{n}_p}^{n_p-1}$  exhibit negligible variation over the dataset  $\mathcal{D}$ .

Denoting  $(A_t, B_t) = (A(p(x_t, u_t)), B(p(x_t, u_t)))$  and using the change of variables  $\tilde{b}_i = e^{\tilde{b}_i^L}$ , we aim to achieve (41) under Assumption 5 by solving

$$\min_{\tilde{A}_i, \tilde{B}_i, \tilde{b}_i > 0} \sum_{t=0}^{N-1} \|(A_t, B_t) - (\tilde{A}(\tilde{p}(\tilde{x}_t, u_t)), \tilde{B}(\tilde{p}(\tilde{x}_t, u_t)))\|_2^2, \quad (42)$$

with the biases recovered as  $\tilde{b}_i^L = \log(\tilde{b}_i)$  from the optimizer. Unfortunately, Problem (42) can be as difficult to solve as a system identification problem because of the nonlinear dynamics in (38). To ameliorate this difficulty, we assume that

$$W_i^L \hat{\mathcal{N}}_i(x_t, u_t) \approx W_i^L \hat{\mathcal{N}}_i(\tilde{x}_t, u_t), \quad (43)$$

for  $i \in \mathbb{I}_1^{\tilde{n}_p-1}$  and  $t \in \mathbb{I}_0^{N-1}$ . Note that (43) might be achieved either because  $W_i^L \approx 0$ , or the parameters  $\Theta$  of (14) along with structure of the FNNs. Either way, defining

$$(q_{it}, \tilde{b}_i) := \begin{cases} (e^{W_i^L \hat{\mathcal{N}}_i(x_t, u_t)}, e^{\tilde{b}_i^L}), & i \in \mathbb{I}_1^{\tilde{n}_p-1} \\ (1, 1), & i = \tilde{n}_p, \end{cases} \quad (44)$$

for  $t \in \mathbb{I}_0^{N-1}$ , we approximate (42) as

$$\min_{\tilde{A}_i, \tilde{B}_i, \tilde{b}_i > 0} \sum_{t=0}^{N-1} \left\| \frac{\sum_{j=1}^{\tilde{n}_p} q_{jt} \tilde{b}_j (A_t, B_t) - \sum_{i=1}^{\tilde{n}_p} q_{it} \tilde{b}_i (\tilde{A}_i, \tilde{B}_i)}{\sum_{j=1}^{\tilde{n}_p} q_{jt} \tilde{b}_j} \right\|_2^2 \quad (45)$$

in which we exploit (43) to replace the components  $e^{W_i^L \hat{\mathcal{N}}_i(\tilde{x}_t, u_t)}$  in (40) with their data-based approximations  $e^{W_i^L \hat{\mathcal{N}}_i(x_t, u_t)}$ . While Problem (45) is still a nonlinear programming problem, we first note that the denominator in the objective is larger than 1 because of (44), specifically at  $j = \tilde{n}_p$ . Hence, we can eliminate it from the objective to minimize an upper bound. Secondly, using the change of variables

$$(\hat{A}_i, \hat{B}_i) := \tilde{b}_i (\tilde{A}_i, \tilde{B}_i), \quad i \in \mathbb{I}_1^{\tilde{n}_p}, \quad (46)$$

Problem (45) can be approximated by the QP

$$\min_{\hat{A}_i, \hat{B}_i, \tilde{b}_i > 0} \sum_{t=0}^{N-1} \left\| \sum_{j=1}^{\tilde{n}_p} q_{jt} \tilde{b}_j (A_t, B_t) - \sum_{i=1}^{\tilde{n}_p} q_{it} \hat{A}_i, \hat{B}_i \right\|_2^2. \quad (47)$$

Hence, we propose to solve Problem (47), and recover the system matrices  $(\tilde{A}_i, \tilde{B}_i)$  as in (46) and biases as  $\tilde{b}_i^l = \log(\tilde{b}_i)$ . Then, using the same initial state, we simulate the dynamics in (38), and estimate matrix  $\tilde{C}$  by solving the QP

$$\min_{\tilde{C}} \sum_{t=0}^{N-1} \|y_t - \tilde{C}\tilde{x}_t\|_2^2, \quad (48)$$

where  $\tilde{x}_t$  is the state of (38) with the newly estimated matrices using inputs in the dataset  $\mathcal{D}$ , and  $y_t$  the corresponding plant outputs in the same dataset. Thus, by solving the QPs in (47) and (48), we identify a qLPV system with scheduling variable order  $\tilde{n}_p < n_p$  to approximate (14).

*Remark 2:* While Problem (47) is derived for a single dataset  $\mathcal{D}$ , it can straightforwardly be extended to accommodate multiple datasets.

*Remark 3:* Scheduling variable order reduction can also be achieved by observing the variance of matrices  $(A_i, B_i)$ . If two such pairs are the same, then they can be lumped together. The development of such techniques is a subject of future study.

### B. Parameterization of the RCI set

We present an approach to compute a matrix  $F$  given initial model parameters  $\Theta$  such that the set  $\mathbb{Q}(\Theta)$  in (27) is nonempty. We note that these parameters  $\Theta$  are not expected to capture the plant dynamics well. Rather, they are used to compute a matrix  $F$  that characterizes the RCI set  $\mathbb{X}(q)$ , and will be refined later by solving Problem (32). This is a standard approach in many data-driven robust control synthesis procedures, e.g., [22], [38], etc. For given  $\Theta$ , there exist many techniques to compute an initial matrix  $F$ , e.g., [39]. The approach we adopt, however, explicitly permits the user to trade-off complexity for conservativeness.

Given  $\Theta$ , we first compute the disturbance set parameters  $(c_w, \epsilon_w)$  as in (24). We remark that since  $\Theta$  is expected to induce a large prediction error, the set  $\mathbb{W}_\kappa(c_w, \epsilon_w)$  might be large. Then, we assume to be given a matrix  $\tilde{F} \in \mathbb{R}^{f \times n_x}$  that characterizes a polytope with  $v$  vertices as

$$\{x \in \mathbb{R}^{n_x} : \tilde{F}x \leq 1\} = \text{CH}\{\tilde{x}_l, l \in \mathbb{I}_v^y\}, \quad (49)$$

Typically,  $\tilde{F}$  is chosen such that  $\{x \in \mathbb{R}^{n_x} : \tilde{F}x \leq 1\}$  is the  $\infty$ -norm box in  $\mathbb{R}^{n_x}$  with  $f = 2n_x$  and  $v = 2^{n_x}$ . The following result from [40] can be used to compute a matrix  $F$  characterizing the RCI set  $\mathbb{X}(q)$  for the uncertain LTI system in (21), and consequently the uncertain qLPV system in (20) as per Proposition 3.

*Proposition 7:* Suppose there exists an invertible matrix  $\Sigma \in \mathbb{R}^{n_x \times n_x}$  and vertex control inputs  $u := (u_1, \dots, u_v)$  satisfying the nonlinear inequalities

$$\tilde{F}\Sigma^{-1}(A_i\Sigma\tilde{x}_l + B_i u_l + K_i c_w) + \kappa|F\Sigma^{-1}K_i|\epsilon_w \leq 1, \quad (50a)$$

$$H^y(C\Sigma\tilde{x}_l + c_w) + \kappa|H^y|\epsilon_w \leq h^y, \quad (50b)$$

$$H^u u_l \leq h^u, \quad (50c)$$

for all  $i \in \mathbb{I}_1^{n_p}$  and  $l \in \mathbb{I}_1^v$ . Then, the set  $\mathbb{X}(1)$  defined with  $F = \tilde{F}\Sigma^{-1}$  is RCI for the uncertain LTI system (21).

Following Proposition 7 and recalling that our goal is to synthesize controllers that track output reference signals of

### Algorithm 1 Procedure to solve Problem (30)

**Require:** Dataset  $\mathcal{D}$  and  $\mathcal{D}^w$ , robustness parameter  $\kappa$

- 1: Solve Problem (34) to estimate  $n_p$ , label optimal model parameters as  $\Theta^1$ .
- 2: (Optional) Solve Problem (47) to reduce  $n_p$  further.
- 3: Solve Problem (52) with relaxed constraints, initialized at  $\Theta^1$ . Label optimal model parameters as  $\Theta^2$ .
- 4: Solve Problem (51) with  $\Theta = \Theta^2$  to compute  $F$ .
- 5: Solve Problem (32) initialized at  $\Theta^2$ . Label optimal model parameters as  $\Theta^3$ .
- 6: If  $\mathbb{Q}(\Theta^3)$  is nonempty, return  $\Theta^3$ . Else, re-solve Problem (32) with larger  $\tau_c$ .

the form  $\{y_k^r \in \mathbb{R}^{n_y}, k \in \mathbb{I}_0^M\}$ , we propose to compute  $\Sigma$  by solving the nonlinear programming problem

$$\min_{\Sigma, u, v} \sum_{k=0}^M \|y_k^r - C z_k\|_2^2 \quad (51)$$

$$\text{s.t.} \begin{cases} (50), z_0 = 0, v_k \in \mathbb{U}, \tilde{F}\Sigma^{-1}z_k \leq 1, \\ z_{k+1} = A(p(z_k, v_k))z_k + B(p(z_k, v_k))v_k, k \in \mathbb{I}_0^{M-1}, \end{cases}$$

formulated in a similar way as (29). We remark that Problem (51) simplifies significantly if the initial model is LTI, i.e.,  $(A_i, B_i, K_i) = (\bar{A}, \bar{B}, \bar{K})$  for all  $i$ . In such cases, the constraints in (50) need only be enforced for a single system, and the dynamics constraint in (51) reduces to  $\bar{x}^+ = \bar{A}\bar{x} + \bar{B}\bar{u}$ . Assuming that (51) is feasible, the matrix  $F$  is computed as  $F = \tilde{F}\Sigma^{-1}$ , following which the configuration-constraint matrix  $E$  and vertex maps  $\{V_l, l \in \mathbb{I}_1^v\}$  are computed following [17, Theorem 2].

Given matrix  $\tilde{F}$  satisfying (49), we now present a heuristic approach to compute system parameters  $\Theta$  that encourage feasibility of Problem (51). We attempt to compute model parameters  $\Theta$  such that  $\Sigma = I_{n_x}$ , i.e., the identity matrix, satisfies (50). To this end, we formulate the problem

$$\min_{\Theta, x_0, u} \frac{1}{N} \sum_{t=0}^{N-1} \|y_t - Cx_t\|_2^2 \quad (52)$$

$$\text{s.t.} \begin{cases} x_{t+1} = A(p(x_t, u_t))x_t + B(p(x_t, u_t))u_t, \\ \tilde{F}(A_i\tilde{x}_l + B_i u_l + K_i c_w) + \kappa|FK_i|\epsilon_w \leq 1, \\ H^y(C\Sigma\tilde{x}_l + c_w) + \kappa|H^y|\epsilon_w \leq h^y, \\ H^u u_l \leq h^u, i \in \mathbb{I}_1^{n_p}, l \in \mathbb{I}_1^v, t \in \mathbb{I}_1^{N-1}. \end{cases}$$

Note that Problem (52) is formulated by posing the conditions in (50) formulated with  $\Sigma = I_{n_x}$  as constraints. To solve it, we relax the inequality constraints with a large penalty parameter. In practice, this approach was observed to compute models  $\Theta$  that render Problem (51) feasible.

### C. qLPV SysID with RCI regularization

In Algorithm 1, we outline the procedure to formulate and solve Problem (30). Firstly, we follow the procedure in Section IV-A to identify the scheduling variable order  $n_p$ . Then, we solve Problem (52) with relaxed constraints, using

a solver initialized with the solution of the previous step. We use the computed model parameters  $\Theta^2$  to solve the nonlinear programming problem (51) to compute  $F$ , using which we compute matrices  $E$  and  $V$  required to formulate the conditions in (22) and characterize the set  $\mathcal{Q}(\Theta)$ . Note that if Problem (51) is feasible with  $\Theta = \Theta^2$ , then  $\mathcal{Q}(\Theta^2)$  is nonempty. Finally, we solve Problem (32), with the solver initialized at  $\Theta^2$ . We choose the penalty parameter  $\tau_c$  to be large enough that the constraints of Problem (30) are satisfied at the solution, or equivalently, the set  $\mathcal{Q}(\Theta^3)$  is nonempty thus solving Problem 1.

#### D. Challenges and limitations

We identify the following challenges in the implementation of Algorithm 1.

1) *Trajectory Length*: Using datasets  $\mathcal{D}$  with long trajectories (large  $N$ ) can make the identification problem difficult to solve due to the instability of the dynamics during optimization [41]. For our qLPV parameterization in (14), a reasonable heuristic to mitigate this involves first identifying a stable LTI system  $x^+ = Ax + Bu$  that best fits the data, and initializing  $(A_i, B_i) = (A, B)$  for all  $i \in \mathbb{I}_1^{n_p}$  such that the qLPV system behaves as a stable LTI system for all  $p \in \mathcal{P}$ .

2) *Robustness Parameter*: Selecting the disturbance set parameter  $\kappa$  involves a trade-off. As discussed previously, while larger values yield more robust models, they may render the RCI set synthesis infeasible.

3) *Initial RCI Set*: While the model identified by solving Problem (52) with relaxed constraints was generally observed to result in feasibility of Problem (51) to compute the matrix  $F$ , this is not guaranteed. Furthermore, Problem (52) might fail simply due to unstable and uncontrollable modes in the data, rendering the original Problem (1) inherently infeasible. Distinguishing between numerical infeasibility and this fundamental structural infeasibility remains a subject for future study. From a numerical standpoint, feasibility can be improved by choosing a more expressive template matrix  $\tilde{F}$ . However, this may result in a polytope  $\mathbb{X}(q)$  with a large number of vertices, significantly increasing the complexity of the set  $\mathcal{Q}(\Theta)$  and Problem (30). An alternative is to co-synthesize a linear feedback controller for the uncertain LTI system in (21) to obtain a low-complexity RCI set, though this typically reduces control authority.

4) *Scalability*: Since our approach relies on polytopic RCI sets, scalability with the state dimension  $n_x$  remains a concern, since the complexity of a polytope in terms of number of facets and vertices can increase exponentially with the dimension. An approach to tackle this complexity is to carefully compute the template matrix  $F$  using a structured approach, such as that presented in [42].

## V. NUMERICAL RESULTS

### A. Effectiveness of the qLPV parameterization

We want to illustrate the effectiveness of the proposed parameterizations for system identification and control synthesis through numerical examples. First, we want to show

the effectiveness of the qLPV parameterization, in which we solve Problem (30) with the constraints ignored, i.e.,

$$\begin{aligned} \min_{\Theta, x_0} \quad & \frac{1}{N} \sum_{t=0}^{N-1} \|y_t - Cx_t\|_2^2 \\ \text{s.t.} \quad & x_{t+1} = A(p(x_t, u_t))x_t + B(p(x_t, u_t))u_t, \quad t \in \mathbb{I}_0^{N-1}. \end{aligned} \quad (53)$$

We solve Problem (53) using the `jax-sysid` library [35] with scaled data  $u \leftarrow (u - \mu_u)/\sigma_u$  and  $y \leftarrow (y - \mu_y)/\sigma_y$ , where  $\mu_u, \mu_y$  and  $\sigma_u, \sigma_y$  are the empirical mean and standard deviation of the inputs and outputs in the Training dataset. The quality of fit is measured using the best fit ratio (BFR)

$$\text{BFR} := \max \left\{ 0, 1 - \sqrt{\frac{\|y_t - \hat{y}_t\|_2^2}{\|y_t - (1/N) \sum_{t=0}^{N-1} y_t\|_2^2}} \right\} \times 100,$$

where  $y_t$  and  $\hat{y}_t$  are the measured and predicted outputs respectively. When evaluating BFR, we use the entire dataset to estimate the initial state of the qLPV model. For all examples, we solve Problem (53) by first running 2000 Adam iterations [33], followed by a maximum of 5000 iterations of the L-BFGS-B solver [34]. We start the optimization with a randomly initialized  $\Theta$ <sup>1</sup>.

1) *Trigonometric system*: We consider the problem of identifying a qLPV model of the nonlinear plant

$$\begin{aligned} z^+ &= \begin{bmatrix} a_1 \sin(z_1) + b_1 \cos(0.5z_2)u \\ a_2 \sin(z_1 + z_3) + b_2 \arctan(z_1 + z_2) \\ a_3 e^{-z_2} + b_3 \sin(-0.5z_1)u \end{bmatrix} + q_x, \\ y &= \arctan(c_1 z_1^3) + \arctan(c_2 z_2^3) + \arctan(c_3 z_3^3) + q_y \end{aligned}$$

where the additive noise terms  $q_x \sim 0.01\mathcal{N}^3$ ,  $q_y \sim 0.01\mathcal{N}$ , and  $a = [0.5, 0.6, 0.4]$ ,  $b = [1.7, 0.4, 0.9]$ ,  $c = [2.2, 1.8, -1]$ . The training and test datasets, consisting of 5000 points each, are built by exciting the system using inputs  $u$  uniformly distributed in  $[-0.5, 0.5]$ . To identify a qLPV model, we solve Problem (53) with  $n_x = 3$ , scheduling variable dimension  $n_p \in \{1, 2, 3, 4\}$ , and number of hidden layers  $n_h \in \{1, 2\}$  for each FNN  $\mathcal{N}_i(x, u)$  defining the scheduling function as in (16). We use 6 swish activation units in each layer. The obtained BFR scores for  $n_p = 1$ , which corresponds to identifying an LTI system with  $n_x = 3$  are 67.962 over the training set and 68.449 over the test set, indicating a poor fit quality. For qLPV identification, the BFR scores are shown in Table I. We observe that using the proposed qLPV parameterization improves the fit quality.

In order to study the effect of group Lasso regularization in (34) for reducing the scheduling variable order, we solve Problem (34) with  $n_x = 2$ ,  $n_p = 10$ ,  $n_h = 1$ , and the FNN  $\mathcal{N}_i(x, u)$  parameterized with 3 swish activation units. In Figure 2, we plot the BFR scores and reduced scheduling variable order as  $\kappa_p$  increases. We report that  $\tilde{n}_p = 10$  for all  $\kappa_p \in [0, 10^{-6}]$ . For  $\kappa_p \geq 2.75 \times 10^{-6}$ , we obtain  $\tilde{n}_p = 1$ , such that the qLPV system reduces to an LTI system.

<sup>1</sup>Corresponding code found on [https://github.com/samku/qLPV\\_concurrent\\_identification](https://github.com/samku/qLPV_concurrent_identification)

$n_p$	$n_h = 1$		$n_h = 2$	
	Training	Test	Training	Test
2	93.426	94.060	94.259	93.262
3	95.154	96.022	95.995	95.223
4	95.978	96.331	96.120	96.143

TABLE I: BFR scores for the trigonometric system with  $n_x = 3$ .

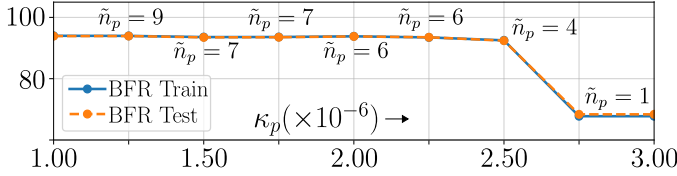


Fig. 2: Variation of BFR with group Lasso regularization parameter  $\kappa_p$  in Problem (34) with  $n_p = 10$ , along with reduced scheduling variable  $\tilde{n}_p$  order following Proposition 6. We obtain  $\tilde{n}_p = 10$  for  $\kappa_p \in [0, 10^{-6}]$  and  $\tilde{n}_p = 1$  for  $\kappa_p \geq 2.75 \times 10^{-6}$ .

2) **Benchmark SysID examples:** We solve Problem (53) using datasets generated by benchmark systems. The BFR scores are shown in Table II. In all examples, we set  $n_x = 4$  and  $n_p = 3$ , and parameterize each FNN  $\mathcal{N}_i(x, u)$  with a single hidden later containing 6 swish activation units. Further, we append the regularization  $1e^{-4} \|\Theta\|_2^2$  to improve generalization quality of the identified model. The datasets are indicated as

- [2T] Two-tank system from [43], with 2300 training and 700 test data.
- [SB] Silverbox system from [44], with 7950 training and 1023 test data.
- [HW] Hammerstein-Wiener system from [45], with 75000 training and 108000 test data.
- [MR] Magneto-Rheological Fluid Damper system from [46], with 2000 training and 1499 test data.

In Table II, we also indicate the BFR scores when identifying an LTI model, i.e., with  $n_p = 1$ . Observe that the qLPV parameterization achieves a good predictive performance.

For the two-tank system (2T), we now illustrate the a posteriori scheduling order reduction procedure proposed in Section

	$n_p = 1$		$n_p = 3$	
	Training	Test	Training	Test
2T	78.820	83.336	97.152	96.709
SB	78.904	50.201	98.690	98.419
HW	82.449	82.111	95.939	95.089
MR	56.231	49.771	92.717	91.464

TABLE II: BFR scores for the considered benchmark examples.

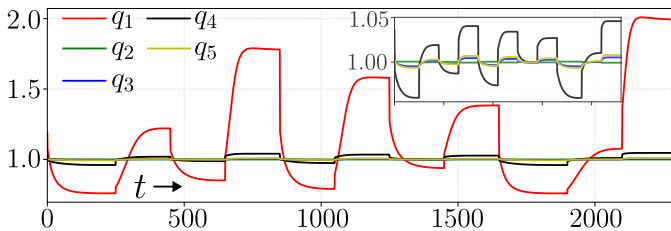


Fig. 3: Trajectories of  $q_{it}$  defined in (44) for the Two-tank system simulated with the training dataset. Observe negligible variation in all except the trajectories of  $q_1$ , thus verifying Assumption 5. The smaller plot is zoomed in on the trajectories of the remaining components.

$\iota$	Training	Test	$\iota$	Training	Test
{2}	94.196	92.599	{2, 3}	94.197	92.586
{2, 3, 5}	94.160	92.741	{2, 3, 5, 4}	88.930	90.005

TABLE III: BFR values for reduced order qLPV model.

IV-A. We consider the over-parameterized qLPV model with  $n_x = 4$  and  $n_p = 6$ , with each FNN parameterized with a single hidden layer containing 6 swish activation units. With this model, we obtain BFR scores of 94.194 and 92.625 over the training and test sets respectively. Towards reducing the scheduling variable order of this model, we plot the values of  $q_{it}$  defined in (44) when simulated using the dataset  $\mathcal{D}$  in Figure 3. We observe that  $q_1$  exhibits the largest variation over the dataset, with the others being almost constant thus verifying Assumption 5. Hence, we eliminate those that are almost constant and estimate the parameters of the reduced-order model in (38) by solving the QPs (47) and (48). In Table III, we see the BFR scores of the reduced order models, with  $\iota$  denoting the eliminated indices. As expected, upto 3 orders can be eliminated without any degradation in the model quality.

### B. SysID with controller synthesis guarantees

We consider data generated by the nonlinear spring-mass-damper system with dynamics

$$\begin{aligned} m_{[1]} \ddot{x}_{[1]} &= 10u - k^s(x_{[1]}) - k^d(\dot{x}_{[1]}) - k^s(\delta x) - k^d(\dot{\delta x}), \\ m_{[2]} \ddot{x}_{[2]} &= -k^s(x_{[2]}) - k^d(x_{[2]}) + k^s(\delta x) + k^d(\dot{\delta x}), \end{aligned}$$

where  $(x_{[1]}, x_{[2]})$  are the positions of the masses. The input  $u$  is the force applied on the first mass, and the output is the position of the second mass. We denote the relative position of the masses as  $\delta x = x_{[1]} - x_{[2]}$ , and the spring and damper reaction forces are  $k^s(x) = ax + bx^3$  and  $k^d(v) = dv + e \cdot \tanh(v/v_0)$  respectively. The plant parameters are  $(m_{[1]}, m_{[2]}, a, b, d, e, v_0) = (0.25, 0.1, 1, 1, 0.5, 0.5, 0.01)$  in appropriate units. The system is simulated from the origin using the Runge-Kutta integrator (Tsit5) implemented in the Diffrax library [47], with a time-step of 0.02s using inputs  $u \in \mathbb{U} = [-1, 1]$ . We excite the system using a combination of multisine and random inputs in this range, using which we build the training dataset  $\mathcal{D}$ , observer dataset  $\mathcal{D}^w$  and test dataset, each with 5000 points. For all these sets, we initialize the plant at the origin. Using the outputs reached by the plant within these datasets, we define the output constraint set as  $\mathbb{Y} = [-1.645, 2.071]$ . We note that the system is nonlinear, indicated by the fact that the best identified LTI system with  $n_x = 4$  achieves BFR scores of 77.131, 70.812 and 70.206 on the training, observer and test datasets respectively. For this system, we aim to design an output reference tracking controller developed using a qLPV model with reduced state dimension  $n_x = 3$ . To this end, we perform the steps outlined in Algorithm 1. We first identify that  $n_p = 3$  is a suitable model order. Then, we parameterize each FNN  $\mathcal{N}_i(x, u)$  in (16) with a single hidden layer and 4 swish activation units. To identify an initial model  $\Theta^1$ , we solve Problem (52) with relaxed constraints and  $\kappa = 1.1$ , in which we select  $\tilde{F} = [I_3 \quad -I_3]^\top$ . In this problem, we set the observer gains

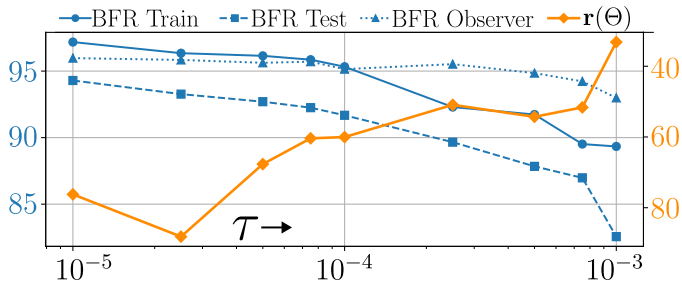


Fig. 4: Variation of BFR and RCI set size  $r(\mathbf{x})$  with  $\tau$ . Observe a tradeoff as  $\tau$  varies. We select  $\tau = 10^{-4}$  for controller synthesis.

$K_i = 0$ , such that the disturbance set is identified purely on the output. While larger values of  $\kappa$  might be selected to increase robustness,  $\kappa = 1.1$  was found sufficient for our simulations. Then, we solve Problem (51) using the IPOPT solver [48] interfaced through CasADi [49], in which we set  $M = 50$ , and formulate it with two reference trajectories  $y_k^r = 2.071$  and  $y_k^r = -1.645$  for all  $k \in \mathbb{I}_0^M$ . We compute an RCI set in which our system can be driven from the origin to the vertices of  $\mathbb{Y}$ . The resulting matrix  $F$  that parameterizes the RCI set  $\mathbb{X}(q)$  is given as  $F = \tilde{F}\Sigma^{-1}$ , where

$$\Sigma^{-1} = \begin{bmatrix} 0.3098 & 2.7690 & 3.0873 \\ 0.1937 & 0.5118 & -0.1277 \\ -0.9419 & 0.3280 & 0.6171 \end{bmatrix}$$

This results in  $\mathbb{X}(q)$  having  $\mathbf{f} = 6$  facets and  $\mathbf{v} = 8$  vertices. For  $\mathbb{X}(1)$ , we compute the matrices  $E$  and  $V$  following the procedure in [17], using which we formulate Problem (32) with penalty parameter  $\tau_c = 1000$ . This penalty was found to be large enough for constraint satisfaction. In Figure 4, we plot the variation in BFR scores, along with the value of the regularization function  $r(\Theta)$  at the solution of Problem (32) for different values of  $\tau$ , with smaller values of  $r(\Theta)$  indicating a less conservative RCI set, as stated in (12). In this plot, BFR Observer refers to the BFR score for the closed-loop observer model (23) over the dataset  $\mathcal{D}^w$ . We note that as  $\tau$  increases, conservativeness of the RCI set reduces at the expense of model quality. For  $\tau \geq 10^{-4}$ , we notice that the BFR observer score sustains at a high value. This is because the model tends to compute large values of the observer gains  $\{K_1, K_2, K_3\}$  to minimize the value of  $\epsilon_w$  defined in (24) in order to reduce the disturbance set size. This, however comes at the price of open-loop prediction inaccuracy as seen by the degradation of BFR Train and BFR Test scores. For comparison, we report that the *sequential design* procedure, which involves first identifying a qLPV model by solving Problem (53), and then computing an RCI set, results in BFR scores of 98.175, 88.686 and 94.002 over the training, test and Observer datasets, with a high conservativeness value of  $r(\Theta) = 349.977$ . This result validates the usage of concurrent synthesis to solve Problem 1 with reduced conservativeness.

Using the model identified with  $\tau = 10^{-4}$ , we then design an output feedback controller for the underlying plant. Given the current plant output  $y_t$ , observer state  $z_t$ , and output

reference  $y_t^r$ , we formulate our controller as

$$(u_t^*, q_t^*) := \arg \min_{u, q} \|y_t^r - Cz^+ \|^2 \quad (54)$$

$$\text{s.t.} \begin{cases} z^+ = A(p)z_t + B(p)u + K(p)(y_t - Cz_t), \\ p = p(z_t, u), \quad u \in \mathbb{U}, \\ q \in \mathbb{Q}(\Theta), \quad z^+ \in \mathbb{X}(q), \end{cases}$$

where  $\mathbb{Q}(\Theta)$  is defined in (27). In Problem (54), we seek to find a feasible control input  $u$  and an RCI set  $\mathbb{X}(q)$  in which the subsequent state belongs, while minimizing the distance between subsequent output and the reference. Recall that our model  $\Theta$  is constructed such that  $\mathbb{Q}(\Theta)$  is nonempty. Furthermore, under Assumption 4, we always have  $y_t - Cz_t \in \mathbb{W}_\kappa(c_w, \epsilon_w)$ . Then, according to Proposition 2, if Problem (54) is initialized with  $\hat{x}_0 - z_0 \in \mathcal{E}$  (where  $\hat{x}_0$  is the state of the underlying plant (5)), then the closed-loop system of the observer (23) and the plant is recursively feasible with  $u_t = u_t^*$ . Since the scheduling variable depends on  $u$ , Problem (54) is a nonlinear optimization problem. By parameterizing the FNNs  $\mathcal{N}_i(z, u)$  to be independent of  $u$ , we can obtain a QP formulation instead. Note that since  $\mathbb{Q}(q)$  is nonempty, one can instead design tube-based MPC controllers, e.g., [31] that are also based on solving a QP online. Such a scheme would be endowed with stability guarantees. Alternatively, by replacing the objective with  $\|u_t^r - u\|_2^2$  where  $u_t^r$  is a reference input signal, safety-filter type schemes [50] can be synthesized.

In Figure 5, we plot the closed-loop trajectories with  $u_t = u_t^*$ , where the reference is composed of piecewise constant signals, and both the observer and plant are initialized at the origin – Since the structure of the plant is known, it can reasonably be assumed that  $\hat{x}_0 - z_0 \in \mathcal{E}$  is verified. We observe that the plant output tracks the reference while being included in the output constraint set  $\mathbb{Y}$ . We observe further that  $y_t$  and  $Cz_t$ , where  $z_t$  is the observer state, are very close to each other. This indicates that our choice of disturbance set satisfies Assumptions 3, such that the controller is able to perform safe tracking following Proposition 2. In Figure 6, we plot the set  $\hat{\mathbb{X}}$  computed as the convex hull of the sets  $\mathbb{X}(q_t^*)$  computed by Problem (54) over simulation length of 20s. We report that  $z_t \in \mathbb{X}(q_t^*)$  always holds resulting in  $z_t \in \hat{\mathbb{X}}$ , demonstrating recursive feasibility.

## VI. CONCLUSIONS

By properly defining a control-oriented regularization term into the system identification problem, we presented an approach to guarantee that the identified model is suitable for designing constrained controllers for the plant generating the data. An uncertain model was derived using a state observer, and the regularization function was characterized as the size of the largest RCI set for the uncertain model. By parameterizing the model as a qLPV model and representing the RCI set as a configuration-constrained polytope, we transformed the identification problem into a computable form. The initialization strategy, which includes estimating suitable number of scheduling signals, enables solving the problem to good-quality solutions as numerically demonstrated.

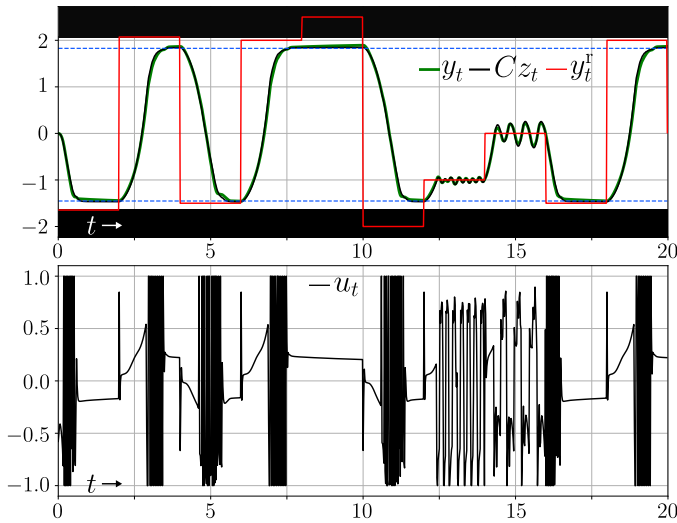


Fig. 5: (Top) Closed-loop trajectories of plant and observer in the output space. Observe that the plant satisfies constraints, indicated by the black regions. The blue line indicates boundaries of the tightened output constraint set  $\mathbb{Y} \ominus \mathbb{W}_{\kappa}(C_w, \epsilon_w)$ ; (Bottom) Corresponding input trajectories generated by (54).

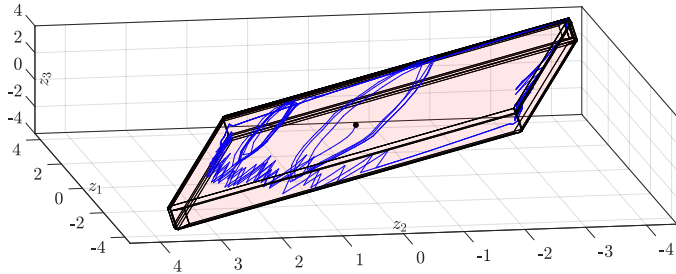


Fig. 6: State trajectories of the closed-loop observer (23), along with the set  $\hat{\mathbb{X}}$  to demonstrate recursive feasibility of Problem (54).

The results of the paper can be extended in several directions, such as: (a) reduce conservativeness in the RCI set by limiting the multiplicative uncertainty encountered when the system evolves within the set; (b) develop alternative formulations of  $r(\Theta)$ ; (c) design tailored optimization algorithms; and (d) integrate the scheme with an active learning framework for data-efficient control-oriented model identification.

## REFERENCES

- [1] P. M. Van Den Hof and R. J. Schrama, "Identification and control — closed-loop issues," *Automatica*, vol. 31, no. 12, pp. 1751–1770, 1995. *Trends in System Identification*.
- [2] R. L. Kosut, M. K. Lau, and S. P. Boyd, "Set-membership identification of systems with parametric and nonparametric uncertainty," *IEEE Transactions on Automatic Control*, vol. 37, no. 7, pp. 929–941, 1992.
- [3] M. Milanese and C. Novara, "Set membership identification of nonlinear systems," *Automatica*, vol. 40, no. 6, pp. 957–975, 2004.
- [4] S. Sadraddini and C. Belta, "Formal guarantees in data-driven model identification and control synthesis," in *Proceedings of the 21st International Conference on Hybrid Systems: Computation and Control (part of CPS Week)*, pp. 147–156, 2018.
- [5] E. Terzi, L. Fagiano, M. Farina, and R. Scattolini, "Learning-based predictive control for linear systems: A unitary approach," *Automatica*, vol. 108, p. 108473, 2019.
- [6] M. Lauricella and L. Fagiano, "Set membership identification of linear systems with guaranteed simulation accuracy," *IEEE Transactions on Automatic Control*, vol. 65, no. 12, pp. 5189–5204, 2020.
- [7] D. Shook, C. Mohtadi, and S. Shah, "A control-relevant identification strategy for GPC," *IEEE Transactions on Automatic Control*, vol. 37, no. 7, pp. 975–980, 1992.
- [8] Y. Chen, H. Peng, J. Grizzle, and N. Ozay, "Data-driven computation of minimal robust control invariant set," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 4052–4058, 2018.
- [9] S. K. Mulagaleti, A. Bemporad, and M. Zanon, "Data-driven synthesis of robust invariant sets and controllers," *IEEE Control Systems Letters*, vol. 6, pp. 1676–1681, 2022.
- [10] M. Zanon and S. Gros, "Safe reinforcement learning using robust MPC," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3638–3652, 2021.
- [11] F. Blanchini and S. Miani, *Set-Theoretic Methods in Control*. Birkhäuser Boston, 2008.
- [12] B. Bamieh and L. Giarre, "Identification of linear parameter varying models," in *Proceedings of the 38th IEEE Conference on Decision and Control*, vol. 2, pp. 1505–1510 vol.2, 1999.
- [13] D. Masti and A. Bemporad, "Learning nonlinear state-space models using autoencoders," *Automatica*, vol. 129, p. 109666, 2021.
- [14] C. Verhoek, G. I. Beintema, S. Haesaert, M. Schoukens, and R. Tóth, "Deep-learning-based identification of LPV models for nonlinear systems," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 3274–3280, 2022.
- [15] G. Beintema, R. Toth, and M. Schoukens, "Nonlinear state-space identification using deep encoder networks," in *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, vol. 144 of *Proceedings of Machine Learning Research*, pp. 241–250, PMLR, 07–08 June 2021.
- [16] S. Formentin and A. Chiuso, "Control-oriented regularization for linear system identification," *Automatica*, vol. 127, p. 109539, 2021.
- [17] M. E. Villanueva, M. A. Müller, and B. Houska, "Configuration-constrained tube MPC," *Automatica*, vol. 163, p. 111543, 2024.
- [18] D. Mayne, S. Raković, R. Findeisen, and F. Allgöwer, "Robust output feedback model predictive control of constrained linear systems," *Automatica*, vol. 42, no. 7, pp. 1217–1222, 2006.
- [19] G. E. G. Kolmanovskiy, Ilya, "Theory and computation of disturbance invariant sets for discrete-time linear systems," *Mathematical Problems in Engineering*, vol. 4, no. 4, pp. 317–367, 1998.
- [20] G. Phanomchoeng and R. Rajamani, "Observer design for lipschitz nonlinear systems using riccati equations," in *Proceedings of the 2010 American Control Conference*, pp. 6060–6065, 2010.
- [21] V. Cerone and D. Regruto, "Set-membership identification of lqv models with uncertain measurements of the time-varying parameter," in *2008 47th IEEE Conference on Decision and Control*, pp. 4491–4496, 2008.
- [22] Y. Chen and N. Ozay, "Data-driven computation of robust control invariant sets with concurrent model selection," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 2, pp. 495–506, 2022.
- [23] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.
- [24] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1702–1717, 2021.
- [25] T. Martin, T. B. Schön, and F. Allgöwer, "Guarantees for data-driven control of nonlinear systems using semidefinite programming: A survey," *Annual Reviews in Control*, vol. 56, p. 100911, 2023.
- [26] C. Verhoek, P. J. W. Koelewyn, S. Haesaert, and R. Tóth, "Direct data-driven state-feedback control of general nonlinear systems," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 3688–3693, 2023.
- [27] R. Tóth, *Modeling and Identification of Linear Parameter-Varying Systems*, vol. 403 of *Lecture Notes in Control and Information Sciences*. Springer, 2010.
- [28] H. D. Nguyen, L. R. Lloyd-Jones, and G. J. McLachlan, "A universal approximation theorem for mixture-of-experts models," *Neural Comput.*, vol. 28, p. 2585–2593, Dec. 2016.
- [29] Q. Zhang, L. Ljung, and R. Pintelon, "On local LTI model coherence for LPV interpolation," *IEEE Transactions on Automatic Control*, vol. 65, no. 8, pp. 3671–3676, 2020.
- [30] P. Trodden, "A one-step approach to computing a polytopic robust positively invariant set," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 4100–4105, 2016.
- [31] F. Badalamenti, S. K. Mulagaleti, A. Bemporad, B. Houska, and M. E. Villanueva, "Configuration-constrained tube MPC for tracking," *IEEE Control Systems Letters*, vol. 8, pp. 1174–1179, 2024.

- [32] J. Nocedal and S. Wright, *Numerical optimization*, pp. 1–664. Springer Series in Operations Research and Financial Engineering, Springer Nature, 2006.
- [33] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015. The paper was initially released on arXiv in 2014.
- [34] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, “A limited memory algorithm for bound constrained optimization,” *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, 1995.
- [35] A. Bemporad, “An L-BFGS-B approach for linear and nonlinear system identification under  $\ell_1$  and group-lasso regularization,” *IEEE Transactions on Automatic Control*, vol. 70, no. 7, pp. 4857–4864, 2025.
- [36] E. J. Olucha, B. Terzin, A. Das, and R. Tóth, “On the reduction of linear parameter-varying state-space models,” 2024.
- [37] P. Koelewijn and R. Tóth, “Scheduling dimension reduction of LPV models - a deep neural network approach,” in *2020 American Control Conference (ACC)*, pp. 1111–1117, 2020.
- [38] M. Mejeri and V. Breschi, “Direct data-driven design of lqv controllers and polytopic invariant sets with cross-covariance noise bounds,” *IEEE Control Systems Letters*, vol. 8, pp. 2427–2432, 2024.
- [39] B. Pluymers, J. Rossiter, J. Suykens, and B. De Moor, “Interpolation based MPC for LPV systems using polyhedral invariant sets,” in *Proceedings of the American Control Conference, 2005*, pp. 810–815 vol. 2, 2005.
- [40] S. K. Mulagaleti, M. Mejeri, and A. Bemporad, “Parameter-dependent robust control invariant sets for LPV systems with bounded parameter-variation rate,” *IEEE Transactions on Automatic Control*, pp. 1–8, 2024.
- [41] A. H. Ribeiro, K. Tiels, J. Umenberger, T. B. Schön, and L. A. Aguirre, “On the smoothness of nonlinear system identification,” *Automatica*, vol. 121, p. 109158, 2020.
- [42] F. Badalamenti, S. K. Mulagaleti, M. E. Villanueva, B. Houska, and A. Bemporad, “Efficient Configuration-Constrained Tube MPC via Variables Restriction and Template Selection,” *arXiv e-prints*, p. arXiv:2505.14440, May 2025.
- [43] MathWorks, “Two-tank system: C mex-file modeling of time-continuous siso system.” MathWorks - MATLAB Documentation, 2023.
- [44] L. Ljung, C. Andersson, K. Tiels, and T. B. Schön, “Deep learning and system identification,” *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 1175–1181, 2020. 21st IFAC World Congress.
- [45] J. Schoukens, M. Vaes, R. Pintelon, and Y. Rolain, “Identification of a wiener-hammerstein benchmark system using time and frequency domain techniques,” in *Proceedings of the 15th IFAC Symposium on System Identification, SYSID*, 2009.
- [46] J. Wang, A. Sano, T. Chen, and B. Huang, “Identification of hammerstein systems without explicit parameterisation of non-linearity,” *International Journal of Control*, vol. 82, p. 937–952, Apr. 2009.
- [47] P. Kidger, *On Neural Differential Equations*. PhD thesis, University of Oxford, 2021.
- [48] A. Wächter and L. T. Biegler, “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, 2006.
- [49] J. Andersson, J. Gillis, G. Horn, J. Rawlings, and M. Diehl, “Casadi – a software framework for nonlinear optimization and optimal control.” <https://web.casadi.org>, 2019.
- [50] K. P. Wabersich and M. N. Zeilinger, “A predictive safety filter for learning-based control of constrained nonlinear dynamical systems,” *Automatica*, vol. 129, p. 109597, 2021.



**Alberto Bemporad** received his M.S. degree cum laude in Electrical Engineering (1993) and his Ph.D. in Control Engineering (1997) from the University of Florence, Italy. He held research positions at Washington University in St. Louis (1996–1997) and ETH Zurich (1997–2002). He served at the University of Siena (1999–2009) and at the University of Trento (2010–2011). Since 2011, he has been Full Professor at the IMT School for Advanced Studies Lucca, where he was Director from 2012 to 2015. He co-founded ODYS S.r.l. in 2011, a company specialized in industrial model predictive control systems, and has held visiting appointments at Stanford University, the University of Michigan, and Zhejiang University. He has authored over 400 publications and 21 patents in model predictive control, hybrid systems, optimization, and automotive control, and co-developed the Model Predictive Control Toolbox (MathWorks) for MATLAB. He was an Associate Editor of the IEEE Transactions on Automatic Control during 2001–2004 and Chair of the IEEE CSS Technical Committee on Hybrid Systems during 2002–2010. He is an IEEE Fellow (since 2010) and IFAC Fellow (since 2025). His awards include the IFAC High-Impact Paper Award (2011–14), IEEE CSS Transition to Practice Award (2019), SAE Environmental Excellence in Transportation Award (2021), the Beale–Orchard-Hays Prize (2024), and an ERC Advanced Grant (2024).



**Sampath Kumar Mulagaleti** received his Bachelor's degree in 2013 from the Amrita School of Engineering (India), Master's degree in 2017 from the KU Leuven (Belgium), both in Mechanical Engineering, and Ph.D. in Systems Science at the IMT School for Advanced Studies Lucca (Italy) in 2023. He held Post-Doc researcher positions at IMT School for Advanced Studies Lucca and University of Trento (Italy) in 2023 and 2024 respectively. He is currently a Post-Doc researcher at IMT School for Advanced Studies Lucca, and is interested in set-based approaches for system identification, analysis and control design.