Risk-Informed Model-Free Safe Control of Linear Parameter-Varying Systems

Babak Esmaeili, *Graduate Student Member, IEEE*, and Hamidreza Modares, *Senior Member, IEEE*

***Abstract*—This paper presents a risk-informed data-driven safe control design approach for a class of stochastic uncertain nonlinear discrete-time systems. The nonlinear system is modeled using linear parameter-varying (LPV) systems. A model-based probabilistic safe controller is first designed to guarantee probabilistic λ-contractivity (i.e., stability and invariance) of the LPV system with respect to a given polyhedral safe set. To obviate the requirement of knowing the LPV system model and to bypass identifying its open-loop model, its closed-loop data-based representation is provided in terms of state and scheduling data as well as a decision variable. It is shown that the variance of the closed-loop system, as well as the probability of safety satisfaction, depends on the decision variable and the noise covariance. A minimum-variance direct data-driven gain-scheduling safe control design approach is presented next by designing the decision variable such that all possible closed-loop system realizations satisfy safety with the highest confidence level. This minimum-variance approach is a control-oriented learning method since it minimizes the variance of the state of the closed-loop system with respect to the safe set, and thus minimizes the risk of safety violation. Unlike the certainty-equivalent approach that results in a risk-neutral control design, the minimum-variance method leads to a risk-averse control design. It is shown that the presented direct risk-averse learning approach requires weaker data richness conditions than existing indirect learning methods based on system identification and can lead to a lower risk of safety violation. Two simulation examples along with an experimental validation on an autonomous vehicle are provided to show the effectiveness of the presented approach.**

***Index Terms*—Data-driven control, linear parameter-varying systems, probabilistic control, safe control.**

I. Introduction

THE past few years have witnessed a surge of attention and advancement in developing safety certificates to equip learning-enabled agents with safety guarantees during and after learning. Specifically, to certify the safety of reinforcement learning (RL) agents, as key enablers of autonomy, various safety certificates have been presented [1]–[7]. These safety certificates leverage control barrier functions (CBFs) to fix their actions myopically [8]–[16]. CBF methods, however, rely on high-fidelity models of the system. As a result, when a system model is not available, an indirect learning approach is typically used to learn a system model from data. However, indirect learning methods that rely on system identification might not be suitable for safety-critical systems with limited available data for the following reasons. First, they can only learn a system model after some richness data conditions on the state-input data are satisfied. Relaxing these data requirements is essential to the success of next-generation safe autonomous systems. Second, the learned open-loop system's variance depends on the signal-to-noise ratio (SNR) of the collected data and cannot be decreased by the control mechanism. Therefore, it is necessary to introduce control-oriented learning methods that minimize variance in safety violations, given current data, to improve safety. Finally, model-based CBF methods for stochastic systems are only limited to the case where the support of noise is finite [11], [17].

Existing results on safe control leverage CBFs to only myopically correct the actions of RL agents when they are not safe. This myopic intervention can lead to convergence to undesired equilibrium [18] and poor performance. Instead, one can learn a safe control policy and merge it with an RL control policy to provide safety and performance guarantees. This will provide a completely model-free paradigm under which both safe and RL controllers are directly learned from data. This is in sharp contrast with CBF methods under which the safety certificate requires the system model, even if the RL controller can be learned directly from data. A challenge is that direct data-driven safe control design for stochastic nonlinear systems is unsettled. Therefore, this paper presented a direct data-driven safe control design method for a class of uncertain stochastic nonlinear systems.

Designing safe controllers that guarantee safety for nonlinear systems is a daunting challenge, even for deterministic systems. This challenge can be overcome by using global linearization methods such as the Koopman operator [19], and local linearization approaches like dynamic-linearization [20], [21]. Another approach is to use linear parameter varying (LPV) systems with compact and convex safe sets by using the Minkowski function [22]. LPV systems are represented by linear systems whose dynamics depend on a set of gain-scheduling parameters that can be estimated or measured during the system's operation. Many nonlinear systems, such as aerospace [23], [24] and various robotic systems [25], [26], can be expressed as LPV systems with unstructured bounded uncertainties or disturbances within a convex set. While learning an explicit model for LPV systems can be data-hungry and

Manuscript received February 23, 2024; revised April 4, 2024; accepted April 14, 2024. This work was supported in part by the Department of Navy award (N00014-22-1-2159) and the National Science Foundation under award (ECCS-2227311). Recommended by Associate Editor Lei Zou. *(Corresponding author: Hamidreza Modares.)*

Citation: B. Esmaeili and H. Modares, "Risk-informed model-free safe control of linear parameter-varying systems," *IEEE/CAA J. Autom. Sinica*, vol. 11, no. 9, pp. 1918–1932, Sept. 2024.

The authors are with the Department of Mechanical Engineering, Michigan State University, East Lansing, MI 48863 USA (e-mail: esmaeil1@msu.edu; modaresh@msu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2024.124479

conservative [27], it is essential to design a safe controller directly from the LPV systems' data without the intermediate step of system identification.

Recently, direct data-driven control has been explored as a way to design safe or optimal control while bypassing the system identification step [28]–[32]. However, the current research on direct data-driven safe control is mainly limited to deterministic systems or treating noise as either a bounded disturbance and designing a robust conservative controller for the system or a measurable signal. Unfortunately, robust control is not effective for systems where noise has a distribution with infinite support. It is of vital importance to design risk-aware data-based safe controllers that account for the variance in safety violations and consequently avoid fluctuations in performance when controllers are used in real systems. As of yet, there is no established method for designing risk-aware safe control for stochastic LPV systems.

This paper introduces a novel risk-informed data-driven approach for designing safe control strategies for stochastic uncertain nonlinear discrete-time systems modeled as LPV systems. A risk-informed control solution is presented for polyhedral safe sets to significantly improve predictability in terms of safety satisfaction and variance reduction of trajectories around the equilibrium point compared to existing risk-neutral methods. To our knowledge, there is no risk-informed safe control solution, even for systems with known dynamics and polyhedral safe sets. The presented approach borrows techniques from set-theoretic control, chance constraints, and primal-dual optimization to design risk-informed controllers relying on the concept of λ-contractivity. A closed-loop data-based presentation of LPV systems is then provided and is leveraged to design control-oriented data-based controllers in which the variance of the closed-loop system is minimized with respect to the safe set. This inherently risk-averse strategy is better suited for safety-critical applications where minimizing the likelihood of safety violations is crucial.

Another significant advantage of the presented approach is its data efficiency. It is shown that one can learn a closed-loop safe controller using a set of data that is not rich enough to learn the open-loop system model from. Developing data-efficient safety certificates is a crucial step toward transitioning from current autonomous agents primarily operating in simulated environments to future agents operating in the physical world. One challenge in this transformation is that safety certificates must be learned and verified using only limited available data collected from the system. This is because, unlike simulated environments under which agents can gain virtually unlimited experience at minimal cost, collecting informative data in high-stakes safety-critical settings might require performing unsafe and costly actions.

The presented method guarantees probabilistic stability and invariance of a given polyhedral safe set. This is in sharp contrast to works such as [28]–[31] that are limited to linear systems and deterministic disturbances, making them overly conservative for systems with stochastic noises. It also differs from CBF-based methods limited to systems with known models with deterministic dynamics [15], [16], [18] or stochastic noises with finite support [11], [17].

The effectiveness of the presented approach is demonstrated through two practical simulation examples, providing practical evidence of its utility and efficacy. In the first example, the devised data-driven safe approach is employed for a magnetic suspension system characterized by parametric uncertainties. The aim is to ensure that the constraints imposed on the position and velocity of the closed-loop suspension system are satisfied. Subsequently, given the critical role of safe control in autonomous vehicles [33]–[35], a practical scenario involving path tracking is explored. The objective is to govern the trajectory of a self-driving car in a manner that avoids violating safety constraints, such as potential collisions with road boundaries. To better reflect real-world conditions, the presence of noise in the environment is taken into account. The risk-aware safe controller presented is then applied to manage the vehicle with assured safety guarantees under these noisy conditions. Additionally, an experimental validation of the approach on an autonomous vehicle is also provided, further substantiating its practical applicability and effectiveness.

Notations: Throughout the paper, the Kronecker product is denoted by $\otimes$ and the Khatri-Rao product, which is a column-wise Kronecker product of two matrices with the same number of columns, is denoted by $\odot$. Moreover, I is the identity matrix with the appropriate dimension and $\mathbf{1}$ is a vector with all of its elements being one. When A is a matrix, A_i refers to its i-th row and A_{ij} is the element in the i-th row and j-th column of A. If A and B are matrices or vectors with the same dimensions, $A(\leq,\geq)B$ denotes a component-wise inequality, i.e., $A_{ij}(\leq,\geq)B_{ij}$ for all i and j. If Q is a matrix and $Q(\leq,\geq)0$, it means that Q is negative or positive semi-definite. Given a set $\mathcal{S}$ and a scalar $\mu \geq 0$, $\mu\mathcal{S}$ is defined as the set of all μx such that $x \in \mathcal{S}$.

The Boolean domain, also known as zero-one-valued domain, is represented by $\mathbb{B}$. A multi-index, which is a collection of indices, is defined as

$$\mathbf{i} = (i_1, \ldots, i_p) \in \mathbb{B}^p \tag{1}$$

where $\mathbb{B}^p = \{\mathbf{i} : i_j \in \mathbb{B}, j = 1, \ldots, p\}$. The set of permutations of the entries of $\mathbf{i}$ is represented as $\mathcal{P}(\mathbf{i})$, and $\mathcal{P}_m(M_{\mathbf{i}})$ denotes a matrix that contains all possible matrices $M_{\mathbf{i}}$ arranged according to the permutations in $\mathcal{P}(\mathbf{i})$.

Also, it is assumed that all random variables are defined on a probability space denoted as $(\Gamma, \mathcal{F}, \mathbb{P})$, where Γ is the sample space, $\mathcal{F}$ is its associated σ-algebra, and $\mathbb{P}$ is the probability measure. Given a random variable $\nu : \Gamma \to \mathbb{R}^n$ defined on the probability space $(\Gamma, \mathcal{F}, \mathbb{P})$, the notation $\nu \in \mathbb{R}^n$ is used to indicate its dimension. The mathematical expectation of ν is denoted as $\mathbb{E}[\nu]$, and if one has $\mathbb{E}[\nu] = \hat{\nu}$, then the covariance of ν can be found using the formula $\mathbb{E}[(\nu - \hat{\nu})(\nu - \hat{\nu})^T]$.

Definition 1 [22]: A C-set is a set that is both convex and compact, and its interior contains the origin.

Definition 2 [22]: A polyhedral C-set, denoted by $\mathcal{S}(F, g)$, is represented by

$$\begin{aligned}\mathcal{S}(F, g) &= \{x \in \mathbb{R}^n : Fx \leq g\} \\ &= \{x \in \mathbb{R}^n : F_j x \leq g_j,\ j = 1, \ldots, q\}\end{aligned} \tag{2}$$

TABLE I
COMPARISON OF EXISTING METHODS AND THE PROPOSED METHODOLOGY

Reference	System type	Disturbance nature	Safety certificate	Data requirements	Key features
[18]	Nonlinear	Deterministic	CBF-based	N/A	Myopic correction of RL agent actions
[19]	Nonlinear	Deterministic	N/A	Moderate	Handles nonlinear systems through global linearization
[20]–[21]	Nonlinear	Deterministic	N/A	Moderate	Handles nonlinear systems through local linearization
[22]	LPV	Deterministic	Minkowski function	Moderate	Compact and convex safe sets for LPV systems
[23]–[26]	LPV	Bounded	N/A	N/A	Handles LPV systems with unstructured uncertainties
[27]	LPV	Bounded	N/A	High	Data-hungry and conservative model learning for LPV systems
[28]–[31]	Linear	Deterministic	Risk-neutral	High	Limited to linear systems and deterministic disturbances, overly conservative
Proposed	LPV	Stochastic	Risk-informed	Low	Handles stochastic LPV systems, probabilistic stability, reduced data requirements

where $F \in \mathbb{R}^{q\times n}$ is a matrix with q rows, i.e., F_j for $j = 1,\ldots, q$, and g is a vector with elements g_j, $j = 1,\ldots,q$.

Lemma 1 [36]: Assume that there is a joint chance constraint denoted by

$$\mathbb{P}[Hx + Mw \le g] \ge (1-\varepsilon) \tag{3}$$

where $x \in \mathbb{R}^n$ represents the decision variable, w is a random variable with a normal distribution $\mathcal{N}(0,\Sigma)$, H and M are matrices with dimensions $q\times n$, and g is a vector in $\mathbb{R}^q$. Now, if the constraints

$$H_j x + M_j \mu \le g_j - k_j \sqrt{M_j \Sigma M_j} \tag{4}$$

are satisfied for all $j = 1,\ldots,q$, where $k_j = \sqrt{\frac{1-\varepsilon_j}{\varepsilon_j}}$ and $\sum_j \varepsilon_j \le \varepsilon$, then the original joint chance constraint (3) is also satisfied.

In Lemma 1, k_j is a constant, and ε_j represents the accepted probability of violation of the constraint $Hx + Mw \le g$.

II. PROBLEM FORMULATION

This section provides a discussion on LPV systems and then states the safe control problem for nonlinear systems that can be modeled by uncertain LPV systems.

A. LPV Systems

In this subsection, LPV and Quasi-LPV systems are discussed, for which a unified safe controller will be designed in the subsequent sections.

LPV systems are a representation of nonlinear systems that offer a structured framework for embedding nonlinearities into varying parameters within a predefined range. LPV surrogate models are often used in practice to describe a large subset of nonlinear systems, thereby providing a valuable tool for solving complex control problems. The LPV systems are typically modeled as

$$x(t+1) = A(\omega(t))x(t) + B(\omega(t))u(t) + w(t) \tag{5}$$

where $x(t) \in \mathbb{R}^n$ denotes the system's state vector, and $u(t) \in \mathbb{R}^m$ represents the control signal. Furthermore, $A(\omega(t))$ and $B(\omega(t))$, referred to as $A(\omega)$ and $B(\omega)$ for simplicity from now on, are the system's internal dynamics and input matrix, respectively. Also, $w(t)$ is the noise vector, and $\omega(t)$ represents the scheduling variables.

Assumption 1: The noise in the system (5), represented by the vector $w(t) = [w_1(t),\ldots,w_n(t)]^T$, has a Gaussian distribution with a mean of zero and variance of Σ, i.e., $w \sim \mathcal{N}(0,\Sigma)$, where $\mathbb{E}[w_i(t)w_j(t)] = 0\ \forall i \ne j$, and $\mathbb{E}[w_i^2(t)] = \sigma_i^2\ \forall i = 1,\ldots, n$.

LPV systems are a versatile class of dynamic systems that effectively capture both linear and nonlinear behavior by considering the system's dynamics as a function of time-varying parameters. These parameters, which encompass physical quantities like operating conditions, inputs, or environmental factors, directly influence the system's behavior. LPV systems offer a flexible framework for modeling and controlling complex systems that display time-varying characteristics. The gain-scheduling variables in LPV systems can be measured in real-time, such as the speed of an aircraft [23], but they cannot be predicted in advance.

LPV systems are typically classified as standard LPV systems, for which the gain-scheduling variables are independent of the systems' states, and quasi-LPV systems, for which the gain-scheduling variables are functions of the systems' states. For LPV systems, the polytopic representation of (5) becomes

$$x(t+1) = \sum_{r=1}^{N_v} \omega_r(t)(A_r x(t) + B_r u(t)) + w(t) \tag{6}$$

where N_v represents the number of vertices, while A_r and B_r refer to the system matrices at the rth vertex, accompanied by its corresponding scheduling variable $\omega_r(t)$.

On the other hand, Quasi-LPV systems relax some of the stringent requirements of LPV systems, enabling more practical modeling and control approaches [37]. In Quasi-LPV modeling, the gain-scheduling variables are expressed as a function of the system states, providing a comprehensive representation that accounts for the interaction between state variables and scheduling parameters.

$$\omega_r(x) = T_r(x(t)),\ r = 1,\ldots,N_v \tag{7}$$

also, $\omega = [\omega_1,\ldots,\omega_r]$ is the vector of gain-scheduling parameters.

This approach enhances the ability to capture the complex dynamics of real-world systems and facilitates the development of accurate and robust control strategies that consider both the system's current state and varying operating condi-

tions.

The matrix components $A(\omega)$ and $B(\omega)$ contain nonlinear state-dependent terms that can be combined and represented as scheduling variables denoted as $T(x) = (T_1(x), \ldots, T_p(x))$. As a result, given that the scheduling variables are continuous and belong to a compact set, there are limits such that $T_j^0 \leq T_j(x) \leq T_j^1$, which allow us to express $T_j(x)$ as $T_j(x) = T_j^0(x)\omega_0^j(x) + T_j^1(x)\omega_1^j(x)$ with $\omega_0^j(x) = (T_j^1 - T_j(x))/(T_j^1 - T_j^0)$ and $\omega_1^j(x) = 1 - \omega_0^j(x)$ for $j = 1, \ldots, p$. Therefore, the system described in (5) can be represented as a polytopic form [38]

$$x(t+1) = \sum_{\mathbf{i} \in \mathbb{B}^p} \omega_{\mathbf{i}}(x)(A_{\mathbf{i}}x(t) + B_{\mathbf{i}}u(t)) + w(t) \tag{8}$$

where $\omega_{\mathbf{i}}(x)$ are the scheduling variables satisfying the convex properties, i.e., $\sum_{\mathbf{i} \in \mathbb{B}^p} \omega_{\mathbf{i}}(x) = 1$ and $\omega_{\mathbf{i}}(x) \geq 0$, and they are computed as follows:

$$\omega_{\mathbf{i}}(x) = \prod_{j=1}^{p} \omega_{i_j}^j(x) \tag{9}$$

with $i_j \in \mathbb{B}$, where $\mathbb{B} = \{0, 1\}$. The subscript $\mathbf{i}$ is a multi-index defined in (1). Also, $A_{\mathbf{i}} \in \mathbb{R}^{n \times n}$ and $B_{\mathbf{i}} \in \mathbb{R}^{n \times m}$ are vertices of the polytope.

Equation (8) can also be written as follows:

$$x(t+1) = A_s(\omega \otimes x(t)) + B_s(\omega \otimes u(t)) + w(t) \tag{10}$$

where $A_s = \mathcal{P}_m(A_{\mathbf{i}})$ and $B_s = \mathcal{P}_m(B_{\mathbf{i}})$.

Assumption 2: The scheduling map $T(x(t))$ and the number of vertices in the polytopes, which is $N_v = 2^p$, are both known.

This assumption is a common and standard assumption for creating gain scheduling maps using prior knowledge of system dynamics, as described in [39]. This prior knowledge typically encompasses information about the system's behavior, dynamics, or other pertinent characteristics. The methods detailed in [39] facilitate the development of scheduling maps that adjust the control parameters of a system based on its operational conditions. These maps are devised with a comprehensive understanding of the underlying system behavior, ensuring that the control modifications are suitable and effective. By assuming the knowledge of the scheduling map $T(x(t))$ and the number of vertices $N_v = 2^p$, our research builds on this proven methodology and capitalizes on the advantages of gain scheduling to enhance control performance.

In this paper, since a direct data-driven control is presented, it is not necessary to have an explicit LPV model, and only the scheduling map is required. That is, it is assumed that the system matrices $A_{\mathbf{i}}$ and $B_{\mathbf{i}}$ are unknown (e.g., due to parametric uncertainties in the original nonlinear system for Quasi-LPV systems). To directly learn a safe controller for its corresponding LPV model, data collected from the original system will be used. It is worth noting that, for the remainder of this paper, the Quasi-LPV modeling will be considered for the control design procedure without loss of generality. Nevertheless, it is important to highlight that the theoretical results also hold true for the LPV modeling.

B. Probabilistic Safe Control Design: Problem Formulation

This subsection provides the problem formulation for probabilistic safe control of stochastic LPV systems under safety constraints.

Before presenting the problem statement, given that set invariance is the primary technique used to ensure safety, the following definitions are provided to clarify the concept. These definitions help to establish a framework for ensuring that the system remains within a predetermined set of states over time, which is essential for safety-critical applications, and to facilitate the design of controllers that can enforce set invariance.

Definition 3 [40]: For the system (5), the set $\mathcal{S}$ is a positive invariant set in probability (ISiP) if the initial state $x(0)$ is in $\mathcal{S}$, then the probability of the state $x(t)$ remaining in the set $\mathcal{S}$ for all $t \geq 0$ is at least $(1 - \varepsilon)$, where ε is a small but acceptable risk level.

To ensure that a safe set is ISiP, the concept of λ-contractivity can be leveraged, which is defined next.

Definition 4: For the system (5), the set $\mathcal{S}$ is considered to be λ-contractive in probability if, for a given value of $0 \leq \lambda < 1$, when the state $x(t)$ is in the set $\mathcal{S}$, the probability of the next state $x(t+1)$ being contained within a scaled version of $\mathcal{S}$ is at least $(1 - \varepsilon)$ for all $t \geq 0$, i.e., $\mathbb{P}[x(t+1) \in \lambda\mathcal{S}] \geq (1 - \varepsilon)$, where ε is a small but acceptable risk level.

The λ-contractive in probability property is important for ensuring that the system remains safe and stable while accounting for the probabilistic nature of the system dynamics. As shown in [32], λ-contractive in probability property guarantees the ISiP. Therefore, by definition of the ISiP, once the system starts from the safe sets, it does not leave it with a high probability, which guarantees probabilistic safety. The contraction rate λ controls the speed of convergence, with larger values leading to faster convergence but potentially higher oscillations and overshoots.

Problem 1: Consider a nonlinear system modeled by the stochastic LPV system (5). Design a gain-scheduling controller in the form of

$$u(t) = K(\omega)x(t) = \sum_{\mathbf{i} \in \mathbb{B}^p} \omega_{\mathbf{i}}(x)K_{\mathbf{i}}x(t) \tag{11}$$

such that a given polyhedral safe set $\mathcal{S}(F, g)$ remains ISiP, or equivalently it satisfies the λ-contractive in probability property.

Both model-based and data-based solutions to this problem will be provided. The following assumption, followed by a related definition is needed to solve this problem.

Definition 5 [22]: The system (5) is called gain scheduling stabilizable, if a controller in the form of $u(t) = K(\omega)x(t)$ exists such that the nominal closed-loop system $x(t+1) = (A(\omega) + B(\omega)K(\omega))x(t)$ is globally asymptotically stable.

Assumption 3: The system (5) is gain-scheduling stabilizable.

Assumptions 2 and 3 are commonly used in the field of LPV systems, as demonstrated in [38] and [41]. The scheduling map $T(x(t))$ and the number of vertices $N_v = 2^p$ can be determined according to the specific nonlinearities present in

the system model.

Now, a closed-loop form of the LPV system under the gain-scheduling control of the form (11) is presented.

Lemma 2: Consider that Assumptions 2 and 3 are satisfied for the system (5). Then, using the controller (11), the closed-loop LPV system can be expressed as

$$x(t+1)=\sum_{\mathbf{i}\in\mathbb{B}^p}\omega_{\mathbf{i}}(x)\sum_{\mathbf{j}\in\mathbb{B}^p}\omega_{\mathbf{j}}(x)(A_{\mathbf{i}}+B_{\mathbf{i}}K_{\mathbf{j}})x(t)+w(t) \tag{12}$$

where $\mathbf{i}=(i_1,\ldots,i_p)$ and $\mathbf{j}=(j_1,\ldots,j_p)$ are multi-indices.

Proof: Substituting the gain-scheduling controller (11) into the LPV system (8) results in

$$x(t+1)=\sum_{\mathbf{i}\in\mathbb{B}^p}(\omega_{\mathbf{i}}(x)A_{\mathbf{i}}x(t))+\sum_{\mathbf{i}\in\mathbb{B}^p}(\omega_{\mathbf{i}}(x)B_{\mathbf{i}})\sum_{\mathbf{j}\in\mathbb{B}^p}(\omega_{\mathbf{j}}(x)K_{\mathbf{j}}x(t))+w(t). \tag{13}$$

Also, due to the fact that $\sum_{\mathbf{j}\in\mathbb{B}^p}\omega_{\mathbf{j}}(x)=1$, (13) can be rewritten as

$$x(t+1)=\sum_{\mathbf{i}\in\mathbb{B}^p}\omega_{\mathbf{i}}(x)(\sum_{\mathbf{j}\in\mathbb{B}^p}\omega_{\mathbf{j}}(x)A_{\mathbf{i}}x(t)+B_{\mathbf{i}}\sum_{\mathbf{j}\in\mathbb{B}^p}\omega_{\mathbf{j}}(x)K_{\mathbf{j}}x(t))+w(t). \tag{14}$$

After some mathematical manipulation, (14) is equivalent to (12), which concludes the proof. ■

Remark 1: It should be mentioned that when the input matrix, i.e., $B(\omega)$, is fixed, the term $\sum_{\mathbf{i}\in\mathbb{B}^p}(\omega_{\mathbf{i}}(x)B_{\mathbf{i}})$ in (13) becomes a constant matrix B, and hence, one gets

$$x(t+1)=\sum_{\mathbf{i}\in\mathbb{B}^p}\omega_{\mathbf{i}}(x)(A_{\mathbf{i}}+BK_{\mathbf{i}})x(t)+w(t). \tag{15}$$

LPV system (12) can be written as the following compact form:

$$x(t+1)=((A_s\otimes\mathbf{1}_{2^p}^T)+(B_s\otimes\mathbf{1}_{2^p}^T)(I_{2^p\times 2^p}\otimes K_{s,d}))\times((\omega\otimes\omega)\otimes x(t))+w(t) \tag{16}$$

where $K_{s,d}$ is a block-diagonal matrix in which each diagonal component is $K_{\mathbf{i}}$ for $\mathbf{i}\in\mathbb{B}^p$. By defining $A_s'=A_s\otimes\mathbf{1}_{2^p}^T$, $B_s'=B_s\otimes\mathbf{1}_{2^p}^T$, and $K_s'=I_{2^p\times 2^p}\otimes K_{s,d}$, (16) can also be represented as

$$x(t+1)=(A_s'+B_s'K_s')((\omega\otimes\omega)\otimes x(t))+w(t). \tag{17}$$

In the literature, there is no existing risk-informed solution for even linear systems with known dynamics and polyhedral safe sets. To address this gap, we design a risk-informed controller by integrating techniques from set-theoretic control, primal-dual optimization, and chance constraint. In the subsequent sections, we present model-based and model-free solutions to Problem 1. More specifically, Theorem 1 in Section III provides a model-based solution to Problem 1. Since the system models are not typically available, Theorem 2 in Section IV-A leverages the results of Theorem 1 to provide a data-based solution. Theorem 2, however, assumes that the noise is measurable, which is not realistic. To provide a risk-informed safe control solution that solves Problem 1 for systems with unmeasurable noise, Theorem 3 is presented in Section IV-B. That is, Theorem 3 extends the results of Theorem 2 for the case where the system noise cannot be measured.

III. Model-Based Design: A Probabilistic Safe Control Approach

This section presents a new solution for designing a model-based controller for Problem 1. To this end, the presented method establishes conditions for λ-contractiveness of the safe set for the system (5). Establishing this condition is challenging since it requires finding a risk-dependent term that tightens the constraints to ensure achieving the desired risk level. We formalize an optimization to characterize the risk-aware or probabilistic set contractiveness and then provide its dual optimization to find this term in a non-conservative manner.

The following theorem outlines the conditions required for the safe set to be λ-contractive, ensuring that the probabilistic behavior of the system remains within a scaled version of the safe set. By satisfying these conditions, the model-based control design can guarantee both safety and stability, even in the presence of noise.

Theorem 1: Consider the LPV system (8) satisfying Assumptions 1–3 with the controller (11). Then, the polyhedral set $\mathcal{S}(F,g)$ is λ-contractive in probability for the closed-loop system if and only if there exist non-negative matrices $P_{\mathbf{i}}\geq 0$, $\mathbf{i}=(i_1,\ldots,i_p)$ such that

$$P_{\mathbf{i}}F=F(A_{\mathbf{i}}+B_{\mathbf{i}}K_{\mathbf{i}}) \tag{18}$$

$$P_{\mathbf{i}}g\leq\lambda g-l \tag{19}$$

where $l=(l_1,\ldots,l_q)^T$ with

$$l_j=\sqrt{\frac{1-\varepsilon_j}{\varepsilon_j}}\sqrt{F_j\Sigma F_j} \tag{20}$$

for $j=1,\ldots,q$.

Proof: To demonstrate that the conditions (18) and (19) are sufficient to ensure λ-contractiveness of the safe set $\mathcal{S}(F,g)$, one must first identify the conditions necessary for $\mathcal{S}(F,g)$ to be λ-contractive. By satisfying these necessary conditions, it can be shown that the safe set is also λ-contractive, as defined by Definition 7. Specifically, if x belongs to $\mathcal{S}(F,g)$, which is defined by $Fx(t)\leq g$, then λ-contractivity in probability of $\mathcal{S}$ can be ensured by satisfying the following inequality:

$$\mathbb{P}[x(t+1)\in\lambda\mathcal{S}]\geq(1-\varepsilon) \tag{21}$$

or

$$\mathbb{P}[F(A(\omega)+B(\omega)K(\omega))x(t)+Fw(t)\leq\lambda g]\geq(1-\varepsilon). \tag{22}$$

This is the same as the following inequality obtained by applying Lemma 1 on the joint chance constraint:

$$F(A(\omega)+B(\omega)K(\omega))x(t)\leq\lambda g-l. \tag{23}$$

In terms of (12) and (23), one has

$$F\sum_{\mathbf{i}\in\mathbb{B}^p}\omega_{\mathbf{i}}(x)\sum_{\mathbf{j}\in\mathbb{B}^p}\omega_{\mathbf{j}}(x)(A_{\mathbf{i}}+B_{\mathbf{i}}K_{\mathbf{j}})x(t)\leq\lambda g-l. \tag{24}$$

The next step is to show that the inequality in (24) holds for

all ω in the set Ω if it holds for each vertex of Ω. The left side of (24), which involves the polytopic term $F(A(\omega)+B(\omega)K(\omega))x(t)$, reaches its maximum value at one of its vertices. Therefore, (24) and consequently, (22) hold if the following inequality holds for all $\mathbf{i} \in \mathbb{B}^p$:

$$F(A_{\mathbf{i}} + B_{\mathbf{i}}K_{\mathbf{i}})x(t) \leq \lambda g - l. \tag{25}$$

It is now shown that satisfaction of the model-based safety conditions (18) and (19) satisfies (25). By virtue of (18), the left hand-side of (25) becomes

$$F(A_{\mathbf{i}} + B_{\mathbf{i}}K_{\mathbf{i}})x(t) = P_{\mathbf{i}}Fx(t). \tag{26}$$

Also, since $x(t) \in \mathcal{S}$, i.e., $Fx(t) \leq g$, and using (19), one gets

$$F(A_{\mathbf{i}} + B_{\mathbf{i}}K_{\mathbf{i}})x(t) = P_{\mathbf{i}}Fx(t) \leq P_{\mathbf{i}}g \leq \lambda g - l \tag{27}$$

which is the same as (25). The proof of sufficiency is now complete. To demonstrate the necessity condition, suppose that the safe set $\mathcal{S}$ is λ-contractive in probability. It will now be shown that both (18) and (19) are fulfilled. If $x(t) \in \mathcal{S}$, then (22) is satisfied due to λ-contractivity in probability. Consider the linear programming problem below:

$$\gamma_{\mathbf{i},j} = \max_{x} \; F_j((A_{\mathbf{i}} + B_{\mathbf{i}}K_{\mathbf{i}})x + w) \tag{28}$$

$$\text{s.t. } Fx \leq g, \; \forall \mathbf{i} \in \mathbb{B}^p, \; j = 1, \ldots, q. \tag{29}$$

As a result of the λ-contractivity condition in probability given by (23), it follows that $\gamma_{\mathbf{i},j} \leq \lambda g$. This leads to the following dual optimization problem of (28):

$$\gamma_{\mathbf{i},j} = \min_{\xi_{\mathbf{i},j}} \; {\xi_{\mathbf{i},j}}^T g + F_j l \tag{30}$$

$$\text{s.t. } {\xi_{\mathbf{i},j}}^T F = F_j(A_{\mathbf{i}} + B_{\mathbf{i}}K_{\mathbf{i}}), \tag{31}$$

$$\xi_{\mathbf{i},j} \geq 0, \; \forall \mathbf{i} \in \mathbb{B}^p, \; j = 1, \ldots, q. \tag{32}$$

The Lagrangian multipliers for the optimization problem given in (28) are denoted by $\xi_{\mathbf{i},j}$. As the primal problem has a feasible solution, the optimal value of the dual problem is equal to the optimal value of the primal problem.

Define $P_{\mathbf{i}}$ as

$$P_{\mathbf{i}} = \begin{bmatrix} {\xi_{\mathbf{i},1}}^T \\ {\xi_{\mathbf{i},2}}^T \\ \vdots \\ {\xi_{\mathbf{i},q}}^T \end{bmatrix} \in \mathbb{R}^{q \times q}. \tag{33}$$

$P_{\mathbf{i}}$ is non-negative due to the non-negativity of $\xi_{\mathbf{i},j}$. Furthermore, since $\gamma_{\mathbf{i},j}$ is less than or equal to λg according to the λ-contractivity in probability property, then for all $\mathbf{i} \in \mathbb{B}^p$ and $j = 1, \ldots, q$, the inequality ${\xi_{\mathbf{i}j}}^T g + F_j l \leq \lambda g$ holds. Equation (31) is equivalent to (18), and using (33), (27) is equivalent to (19). ■

Corollary 1: Consider the LPV system (8) with the control input (11). Then, λ-contractive in probability property of the polyhedral set $\mathcal{S}(F, g)$ is met for the closed-loop system (12) if and only if a non-negative matrix $P_s = \mathcal{P}_m(P_{\mathbf{i}})^T$ exists such that the following conditions are satisfied:

$$P_s F = (I_{2^p \times 2^p} \otimes F)(A_s + B_s K_s)^T \tag{34}$$

$$P_s G \leq \lambda G - L_s \tag{35}$$

where $K_s = \mathcal{P}_m(K_{\mathbf{i}}) \in \mathbb{R}^{m \times (n \times 2^p)}$ with $K_{\mathbf{i}} \in \mathbb{R}^{m \times n}$, $A_s + B_s K_s = \mathcal{P}_m(A_{\mathbf{i}} + B_{\mathbf{i}}K_{\mathbf{i}})$, $G = (\mathbf{1}_{2^p \times 1} \otimes g)$, and $L_s = \mathbf{1}_{2^p \times 1} \otimes l$ with $l = (l_1, ..., l_q)^T$ defined in (20).

Proof: Theorem 1 states that the closed-loop LPV system represented by (8) and controlled using the method defined in (11) will be λ-contractive in probability with respect to the polyhedral set $\mathcal{S}(F, g)$ if and only if certain conditions are met. Specifically, there must exist 2^p non-negative matrices $P_{\mathbf{i}}$ and $K_{\mathbf{i}}$ for $\mathbf{i} \in \mathbb{B}^p$ that satisfy both (18) and (19). These two conditions can be combined into a single equality and a single inequality, respectively, as follows:

$$\mathcal{P}_m(P_{\mathbf{i}})^T F = (I_{2^p \times 2^p} \otimes F)\mathcal{P}_m(A_{\mathbf{i}} + B_{\mathbf{i}}K_{\mathbf{i}})^T \tag{36}$$

$$\mathcal{P}_m(P_{\mathbf{i}})^T G \leq \lambda G - \begin{bmatrix} l \\ \vdots \\ l \end{bmatrix} \tag{37}$$

which yields (34) and (35), respectively. ■

IV. Data-Driven Probabilistic Safe Control Design

This section aims at presenting a data-based version of conditions (18) and (19) that eliminates the need for the system model in the safe controller. To achieve this, let us suppose that an input sequence of $u(0), u(1), \ldots, u(N-1)$ is applied to the system (5), and N state and gain-scheduling variable samples are collected. These samples are then arranged as follows:

$$U_0 = [u(0), u(1), \ldots, u(N-1)] \tag{38}$$

$$X_0 = [x(0), x(1), \ldots, x(N-1)] \tag{39}$$

$$W_0 = [w(0), w(1), \ldots, w(N-1)] \tag{40}$$

$$\Omega_0 = [\omega(0), \omega(1), \ldots, \omega(N-1)] \tag{41}$$

$$X_1 = [x(1), x(2), \ldots, x(N)] \tag{42}$$

$$X_{\omega\omega} = (\Omega_0 \odot \Omega_0) \otimes X_0 \tag{43}$$

$$U_\omega = \Omega_0 \odot U_0. \tag{44}$$

It should be noted that the scheduling map in (41) is dependent on the system's states, which means that the data matrix Ω_0 can be easily obtained from state measurements. Meanwhile, to learn the dynamics of the LPV system described in (5), it is necessary for the data matrix represented by

$$\begin{bmatrix} U_0 \\ X_{\omega\omega} \end{bmatrix} \tag{45}$$

to have a full-row rank. This condition can be satisfied by collecting at least $N \geq (m+1)n2^p + m2^p$ samples from the LPV

system. However, to directly learn a safe controller, it is shown that a significantly smaller number of samples is required, which only needs to ensure that the matrix $X_{\omega\omega}$ has a full row rank.

These data are then used to derive data-based versions of conditions (18) and (19). The resulting conditions can be directly used to design a safe controller without requiring knowledge of the system model.

Assumption 4: The data matrix $X_{\omega\omega}$ in (43) is full-row rank.

Assumption 4 introduces a data richness condition for the design of a data-based safe controller. Rather than being restrictive, this assumption actually eases the data-richness requirement compared to existing indirect data-based controllers, a point that will be further elaborated in Remark 2 later.

A. Data-Based Safe Control: Available Noise Measurements

Theorem 1 provided a model-based solution to Problem 1. This subsection extends these results to the data-based case and outlines the design procedure of an online data-driven safe control, assuming the availability of noise measurements.

Theorem 2: Consider the LPV system (8). Let the open-loop data be collected and arranged as shown in equations (38)–(44). Let Assumptions 1–4 be satisfied. Then, the polyhedral set $\mathcal{S}(F,g)$ is λ-contractive in probability for the closed-loop system if and only if decision variables $G_{K,s}=\mathcal{P}_m(G_{K,\mathbf{i}})$ and $P_{\mathbf{i}}\geq 0$, for $\mathbf{i}\in\mathbb{B}^p$, exist such that

$$P_{\mathbf{i}}F = F(X_1 - W_0)G_{K,\mathbf{i}} \tag{46}$$

$$P_{\mathbf{i}}g \leq \lambda g - l \tag{47}$$

$$X_{\omega\omega}G_{K,s} = I \tag{48}$$

where $l=(l_1,\ldots,l_q)^T$ is defined in Theorem 1. Moreover, the controller gains are computed as $K_s^T\otimes\mathbf{1}_{2^p}^T = U_\omega G_{K,s}$.

Proof: Using the proof of Lemma 2, (10) is written as

$$x(t+1) = (A_s\otimes\mathbf{1}_{2^p}^T)((\omega\otimes\omega)\otimes x(t)) + B_s(\omega\otimes u(t)) + w(t). \tag{49}$$

Assumption 4 states that $X_{\omega\omega}$ has full row rank, which implies that there exists the right inverse matrix $G_{K,s}$ such that $X_{\omega\omega}G_{K,s}=I$. Using this result and LPV system (49) along with the open-loop data given by (38)−(44), one can express

$$X_1 = A_sX_{\omega\omega} + B_sU_\omega + W_0. \tag{50}$$

By multiplying $G_{K,s}$ on both sides of (50), it yields

$$A_s + B_sU_\omega G_{K,s} = (X_1 - W_0)G_{K,s}. \tag{51}$$

According to $K_s^T\otimes\mathbf{1}_{2^p}^T = U_\omega G_{K,s}$, one has $A_s + B_s(K_s^T\otimes\mathbf{1}_{2^p}^T) = (X_1 - W_0)G_{K,s}$. Furthermore, since, in terms of (11), (8) is equal to (12), it can be concluded that the closed-loop system matrix, i.e., $A_s' + B_s'K_s'$, is equal to $(X_1 - W_0)G_{K,s}$, and hence, (17) is written as the following data-based form:

$$x(t+1) = (X_1 - W_0)G_{K,s}((\omega\otimes\omega)\otimes x(t)) + w(t). \tag{52}$$

Therefore, condition (18) becomes (46) for $\mathbf{i}\in\mathbb{B}^p$. The rest of the proof is analogous to that of Theorem 1. ■

The results of Theorem 2 can be trivially obtained under the assumption that the noise is measurable based on the results of Theorem 1 and a data-based closed-loop representation of the system. However, the controller designed based on Theorem 2 has a drawback in that it needs the measurement of noise, which is not practical. Therefore, to overcome this challenge, a data-driven safe controller based on minimum variance is designed in the subsequent part of the paper. It is a daunting challenge to provide risk-informed controllers with safety guarantees when the noise is not measurable. To achieve this goal, we characterized the set of all the next states given a probabilistic characterization of closed-loop systems and leveraged set containment tools to provide a condition under which risk-informed safety is guaranteed. We also provided new optimization formulations to provide minimum-variance controllers.

B. Minimum Variance-Based Method: Unavailable Noise Measurements

In this subsection, a minimum variance-based approach is presented to relax the restrictive assumption regarding the availability of noise measurements. This approach aims to mitigate the limitations associated with this assumption.

When using indirect learning approaches [30], [31], predetermined high-confidence sets are assigned for dynamics $A(\omega)$ and $B(\omega)$. This means that the controller $K(\omega)$ can only impact the variance associated with the $B(\omega)K(\omega)$ portion of the closed-loop dynamics. Conversely, with the designed minimum variance-based direct learning, the entire closed-loop dynamics $A(\omega)+B(\omega)K(\omega)$ is learned, and the control gain $K(\omega)$ can be developed to decrease the variance for the entire dynamics.

To begin with, according to (52), the nominal model of $A_s + B_sU_\omega G_{K,s}$ is $X_1G_{K,s}$. Thereupon, the nominal next state is computed as

$$x_n(t+1) = X_1G_{K,s}((\omega\otimes\omega)\otimes x(t)). \tag{53}$$

Then, the random part of $x(t+1)$ is obtained as

$$\begin{aligned} x_r(t+1) &= x(t+1) - x_n(t+1) \\ &= -W_0G_{K,s}((\omega\otimes\omega)\otimes x(t)) + w(t). \end{aligned} \tag{54}$$

By defining $\bar{x}(t) = ((\omega\otimes\omega)\otimes x(t))$, the variance of $x_r(t+1)$ is computed as

$$\begin{aligned} \mathbb{E}[x_r(t+1)x_r^T(t+1)] &= \mathbb{E}[W_0G_{K,s}\bar{x}(t)\bar{x}^T(t)G_{K,s}^TW_0^T] + \Sigma \\ &= \Sigma\mathrm{Tr}(\bar{x}(t)\bar{x}^T(t)G_{K,s}^TG_{K,s}) + \Sigma. \end{aligned} \tag{55}$$

The last equality in (55) is obtained owing to the fact that for a given random vector $\nu\in\mathbb{R}^{n\times 1}$ and a matrix $Q\in\mathbb{R}^{n\times n}$, one has [42]

$$\mathbb{E}[\nu^TQ\nu] = \mathrm{Tr}(Q\mathbb{E}[\tilde{\nu}\tilde{\nu}^T]) + \mathbb{E}[\nu]^TQ\mathbb{E}[\nu] \tag{56}$$

where $\tilde{\nu} = \nu - \mathbb{E}[\nu]$. Now, based on the Cauchi-Schwarz inequality [43], one gets

$$\begin{aligned}\mathrm{Tr}(\bar{x}(t)\bar{x}^T(t)G_{K,s}^TG_{K,s}) &\le \left|\mathrm{Tr}(\bar{x}(t)\bar{x}^T(t)G_{K,s}^TG_{K,s})\right| \\ &\le \sqrt{\mathrm{Tr}\left((\bar{x}(t)\bar{x}^T(t))^2\right)\mathrm{Tr}((G_{K,s}^TG_{K,s})^2)} \\ &\le \sqrt{\mathrm{Tr}(\bar{x}(t)\bar{x}^T(t))^2\mathrm{Tr}(G_{K,s}^TG_{K,s})^2} \\ &= \bar{x}^T(t)\bar{x}(t)\mathrm{Tr}(G_{K,s}^TG_{K,s}). \end{aligned} \tag{57}$$

Then, it yields

$$\mathbb{E}[x_r(t+1)x_r^T(t+1)] \le \Sigma\|\bar{x}(t)\|^2\mathrm{Tr}(G_{K,s}^TG_{K,s}) + \Sigma. \tag{58}$$

Since $\|\bar{x}(t)\| = \|(\omega\otimes\omega)\otimes x(t)\| = \|\omega\|^2\|x(t)\|$ and $\|\omega\| \le 1$, one gets

$$\begin{aligned}\mathbb{E}[x_r(t+1)x_r^T(t+1)] &\le \Sigma\|\bar{x}(t)\|^2\mathrm{Tr}(G_{K,s}^TG_{K,s}) + \Sigma \\ &\le \Sigma\|x(t)\|^2\mathrm{Tr}(G_{K,s}^TG_{K,s}) + \Sigma. \end{aligned} \tag{59}$$

On account of the inequality $F_jx(t) \le g_j$ for $j = 1,\ldots,q$, it is concluded that

$$\|x(t)\| \le \frac{|g_j|}{\|F_j\|} \Longrightarrow \|x(t)\|^2 \le \Big(\max_j \frac{|g_j|}{\|F_j\|}\Big)^2 = \kappa. \tag{60}$$

Thus, (59) becomes

$$\mathbb{E}[x_r(t+1)x_r^T(t+1)] \le \bar{V} \tag{61}$$

where $\bar{V} = \Sigma(\kappa\mathrm{Tr}(G_{K,s}^TG_{K,s}) + 1)$.

Now, the following theorem summarizes the results for the minimum variance-based safe control design technique. This extends the results of Theorem 2 for the more realistic case for which the noise is not measurable.

Theorem 3: Consider the LPV system (8). Let the input/state data be collected by applying an open-loop control sequence to the system and organized by (38), (39), and (41)–(44). Let Assumptions 1–4 be satisfied. Then, the controller $u(t) = K(\omega)x(t)$ guarantees a probabilistic λ-contractive property with risk level δ for the safe set $\mathcal{S}(F,g)$ if there exist matrices $G_{K,\mathbf{i}}$ and non-negative matrices $P_\mathbf{i} \ge 0$ for $\mathbf{i}\in\mathbb{B}^p$ that satisfy the optimization problem given by

$$\min_{P_\mathbf{i},G_{K,\mathbf{i}},\rho} \ \rho \tag{62}$$

$$\text{s.t. } P_\mathbf{i}F = FX_1G_{K,\mathbf{i}} \tag{63}$$

$$P_\mathbf{i}g \le \lambda g - l'_m \tag{64}$$

$$X_{\omega\omega}G_{K,s} = I \tag{65}$$

$$P_\mathbf{i} \ge 0 \tag{66}$$

$$\mathrm{Tr}(G_{K,s}^TG_{K,s}) \le \rho \tag{67}$$

where $l'_m = (l'_{m,1},\ldots,l'_{m,q})^T$ with $l'_{m,j} = \sqrt{F_jV_mF_j^T}$ for $j = 1,\ldots,q$ and $V_m = \Sigma(\kappa\rho+1)(n+2\sqrt{n\log(\frac{1}{\delta})}+2\log(\frac{1}{\delta}))$. Furthermore, the controller gains are calculated as $K_s^T\otimes\mathbf{1}_{2^p}^T = U_\omega G_{K,s}$.

Proof: It is firstly shown that the next state $x(t+1)$ will, with a probability of $1-\delta$, be contained in a particular confidence ellipsoidal set described as

$$\zeta(V,1) = \{x : (x - X_1G_{K,s}\bar{x}(t))^TV^{-1}(x - X_1G_{K,s}\bar{x}(t)) \le 1\}. \tag{68}$$

According to (61) and since $\mathbb{E}[x_r(t+1)] = 0$, $x_r(t+1)$ is a sub-Gaussian random signal with covariance $\bar{V}$. As a result, it is possible to say with a probability of at least $1-\delta$ that [44]

$$x_r(t+1)^T\bar{V}^{-1}x_r(t+1) \le n + 2\sqrt{n\log(\frac{1}{\delta})} + 2\log(\frac{1}{\delta}). \tag{69}$$

This can also be written as

$$x_r(t+1)^TV^{-1}x_r(t+1) \le 1. \tag{70}$$

If the following condition is met, then probabilistic λ-contractivity with the risk level δ is satisfied:

$$x(t)\in\mathcal{S}(F,g) \Rightarrow \mathbb{P}[x(t+1)\in\mathcal{S}(F,\lambda g)] \ge 1-\delta. \tag{71}$$

To ensure that the right-hand side of (71) holds, it is necessary to make sure that the set of all feasible next states is a part of the safe set $\mathcal{S}$ with a probability of $1-\delta$. In other words, it is required that

$$\begin{aligned}&\{x : (x - X_1G_{K,s}\bar{x}(t))^TV^{-1}(x - X_1G_{K,s}\bar{x}(t))\} \\ &\quad\subseteq \{x : Fx \le \lambda g\}. \end{aligned} \tag{72}$$

Equivalently,

$$\begin{aligned}&1 - (x - X_1G_{K,\mathbf{i}}x(t))^TV^{-1}(x - X_1G_{K,\mathbf{i}}x(t)) \ge 0 \\ &\quad\Rightarrow \lambda g - Fx \ge 0, \ \forall\mathbf{i}\in\mathbb{B}^p. \end{aligned} \tag{73}$$

The application of the S-procedure allows for the identification of certain scalars $\tau_{\mathbf{i},j}$ for which the fulfillment of the subsequent condition, for $\mathbf{i}\in\mathbb{B}^p$ and $j = 1,\ldots,q$, is equal to the fulfillment of (73)

$$\begin{aligned}&\lambda g_j - F_jx - \tau_{\mathbf{i},j} \\ &\quad\times[1 - (x - X_1G_{K,\mathbf{i}}x(t))^TV^{-1}(x - X_1G_{K,\mathbf{i}}x(t))] \ge 0, \\ &\quad\forall x, \ \mathbf{i}\in\mathbb{B}^p, \ j = 1,\ldots,q. \end{aligned} \tag{74}$$

By defining $L_{\mathbf{i},j} = \lambda g_j - F_jx - \tau_{\mathbf{i},j}[1 - (x - X_1G_{K,\mathbf{i}}x(t))^TV^{-1}\times(x - X_1G_{K,\mathbf{i}}x(t))]$, the optimal x that minimizes $L_{\mathbf{i},j}$ is calculated as

$$x = (\frac{1}{2}\tau_{\mathbf{i},j}^{-1}V)F_j^T + X_1G_{K,\mathbf{i}}x(t). \tag{75}$$

Substituting (75) into $L_{\mathbf{i},j}$ yields

$$L_{\mathbf{i},j} = \lambda g_j - \frac{1}{4}\tau_{\mathbf{i},j}^{-1}F_jVF_j^T - \tau_{\mathbf{i},j} - F_jX_1G_{K,\mathbf{i}}x(t) \ge 0. \tag{76}$$

Condition (76) is equivalent to the following optimization problem:

$$\beta_{\mathbf{i},j} = \min_x \ L_{\mathbf{i},j} \ge 0 \tag{77}$$

$$\text{s.t. } Fx \le g, \ \forall\mathbf{i}\in\mathbb{B}^p, \ j = 1,\ldots,q. \tag{78}$$

The dual optimization problem can also be written as follows:

$$\beta_{\mathbf{i},j} = \max_{\eta_{\mathbf{i},j}} \ \lambda g_j - \frac{1}{4}\tau_{\mathbf{i},j}^{-1}F_jVF_j^T - \tau_{\mathbf{i},j} - \eta_{\mathbf{i},j}^Tg \tag{79}$$

$$\text{s.t. } \eta_{\mathbf{i},j}^TF = F_jX_1G_{K,\mathbf{i}} \tag{80}$$

$$\eta_{\mathbf{i},j} \geq 0,\ \forall \mathbf{i} \in \mathbb{B}^p,\ j = 1,\ldots,q \tag{81}$$

where $\eta_{\mathbf{i},j}$ are the Lagrangian multipliers. Since a feasible solution exists for the primal problem, the optimal value of the dual problem is equivalent to the optimal value of the primal problem.

To analyze the feasibility of the dual problem (79), and since $0 \leq \beta_{\mathbf{i},j}$, one has

$$\lambda g_j \geq \frac{1}{4}\tau_{\mathbf{i},j}^{-1} F_j V F_j^T + \tau_{\mathbf{i},j} + \eta_{\mathbf{i},j}^T g. \tag{82}$$

By computing first and second derivatives of the right-hand-side of (82) with respect to $\tau_{\mathbf{i},j}$, it is deduced that $\tau_{\mathbf{i},j}^* = \frac{1}{2}\sqrt{F_j V F_j^T}$ is its maximum value. Hence, by substituting $\tau_{\mathbf{i},j}^*$ into (79), it yields

$$\beta_{\mathbf{i},j} = \max_{\eta_{\mathbf{i},j}}\ \lambda g_j - l'_j - \eta_{\mathbf{i},j}^T g \geq 0 \tag{83}$$

$$\text{s.t. } \eta_{\mathbf{i},j}^T F = F_j X_1 G_{K,\mathbf{i}} \tag{84}$$

$$\eta_{\mathbf{i},j} \geq 0,\ \forall \mathbf{i} \in \mathbb{B}^p,\ j = 1,\ldots,q \tag{85}$$

where $l'_j = \sqrt{F_j V F_j^T}$. Equation (83) is equivalent to minimizing l'_j or in other words, $\mathrm{Tr}(G_{K,s}^T G_{K,s})$ in V. Thus, with respect to condition (67), the optimization problem (62) incorporates upper bounds for V and l', denoted as V_m and l'_m, respectively.

Now, consider

$$P_{\mathbf{i}} = \begin{bmatrix} \eta_{\mathbf{i},1}^T \\ \eta_{\mathbf{i},2}^T \\ \vdots \\ \eta_{\mathbf{i},q}^T \end{bmatrix} \in \mathbb{R}^{q\times q}. \tag{86}$$

Therefore, the proof is completed by using (85) to show that the matrix $P_{\mathbf{i}}$ is non-negative. Also, according to (86) and (84), (63) is obtained, and inequality (82) using $\tau_{\mathbf{i},j}^*$ is shown to be equivalent to (64). Next, the decision variable $\rho \geq 0$ satisfying condition (67) is defined, and the minimization problem (62), which is a linear programming, is obtained. ■

Remark 2: Compared to the indirect data-based safe control design methods that rely on system identification, the presented direct learning control scheme requires a considerably smaller amount of independent data to identify the LPV system. Specifically, to identify the system, $2N_v(n+m)$ independent data points must be collected to learn about the open-loop system dynamics. On the other hand, the presented scheme only requires $2N_v n$ independent data points to learn directly about the closed-loop system that satisfies safety. Reducing the required amount of data is advantageous in practical applications, as it can significantly decrease the time and resources needed to collect and process the necessary data. Therefore, the presented scheme is a more efficient and practical approach to data-based safe control. Moreover, the presented approach's reduced data requirement does not lead to higher computational demands. We use a linear program (LP), known for its computational efficiency. The smaller number of data translates into fewer decision variables, lowering computational needs. Also, reducing the number of data samples required for learning avoids risky exploration to collect more data, which is crucial in safety-critical systems.

Remark 3: If vertices of a polyhedral safe set, which are its extreme points, are activated, the corresponding weights of other dynamics, i.e., $(A_{\mathbf{i}}, B_{\mathbf{j}})$ for $\mathbf{i} \neq \mathbf{j}$, become zero. This implies that for designing a safe controller, only the extreme points of a polyhedral safe set are essential in learning the safe control gains during the simulation. In other words, similar to the model-based case, i.e., Theorem 1, other cases do not contribute to the learning process of the data-based safe gains. This consideration ensures that the hypothesis of $X_{\omega\omega}$ becoming full row rank is satisfied during the implementation step.

Remark 4: In this paper, we have developed two types of safe control methods: a model-based approach and a data-based probabilistic approach. When a comprehensive understanding of the LPV system is accessible, the model-based controller outlined in Theorem 1 can be seamlessly applied. However, in situations where the complete model is not accessible, the data-based version of the designed safe controller, as outlined in accordance with Theorem 3, can be employed.

Remark 5: The safe control design based on λ-contractivity is not myopic; it ensures safety not just for the immediate next step, but for all future time steps. This approach aims to identify a feedback controller that guarantees the invariance of a set, ensuring long-term safety. If a linear controller exists that can achieve this, our method is capable of finding it [22].

V. Simulation and Experimental Validation

In this section, two practical examples are provided to validate the efficiency of the designed approach. First, the designed method is applied to a magnetic suspension system, and then, a safe control of an autonomous vehicle is considered in Example 2.

A. Example 1: Magnetic Suspension System

Magnetic suspension systems are used in many engineering applications, such as high-speed trains, frictionless bearings, vehicle suspension systems, and wind tunnels. These systems are inherently nonlinear and exhibit open-loop instability, necessitating closed-loop control to ensure their states remain within a safe operating range. Fig. 1 depicts the magnetic suspension system under closed-loop control using the presented method. A reference point is established to stabilize the ball at

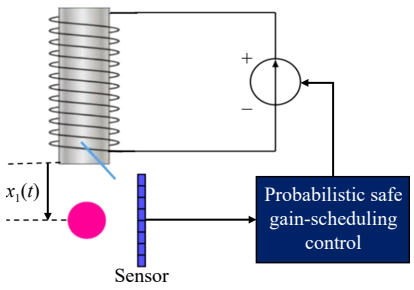

Fig. 1. Control diagram of the magnetic suspension system using the designed safe control approach.

a specific location.

Discrete-time dynamics of a magnetic suspension system is considered as follows [45]:

$$x_1(t+1) = x_1(t) + hx_2(t)$$

$$x_2(t+1) = \frac{hg_a\mu(\mu x_1(t) + 2\mu y_0 + 2)x_1(t)}{(1+\mu(x_1(t)+y_0))^2}x_1(t) + (1 - \frac{hK_m}{m})x_2(t) - \frac{hL\mu}{2m(1+\mu(x_1(t)+y_0))^2}u(t) + w(t). \tag{87}$$

The ball's deviation from its intended position is represented by x_1, while its vertical velocity is represented by x_2. The coil voltage noise is denoted as $w(t)$. The desired position for the ball is set at $y_0 = 0.05$ m. It should be noted that the learning algorithm does not utilize the system dynamics (87), but instead uses it to generate data for learning solely within a simulation environment.

The admissible set for the system states is defined as $C_s = \{x = (x_1, x_2) \in \mathbb{R}^2 \mid |x_1| \le 0.05, |x_2| \le 1\}$. This set is used to establish the boundaries of the scheduling variables. A subset of the admissible set, which is a polyhedral set in the form of (2) with the following F matrix, is considered the safe set:

$$F = \begin{bmatrix} 20.3835 & 1.3397 \\ -20.3835 & -1.3397 \\ -19.8079 & 0.6697 \\ 19.8079 & -0.6697 \end{bmatrix}. \tag{88}$$

The values of the unknown parameters m (the mass of the ball), K_m (the viscosity friction coefficient), and L (the inductance capacity) are not known, but their actual values are $m = 0.068$ kg, $K_m = 0.001$ Nsm^{-1}, and $L = 0.46$ H. The sampling time for the system is set to $h = 0.01$ s, the gravitational constant is $g_a = 9.8$ ms^{-2}, and the coefficient for inductance variation is $\mu = 2$ m^{-1}. Using the Quasi-LPV modeling, the scheduling variables in (7) are defined as follows:

$$T_1(x) = \frac{hg_a\mu(\mu x_1(t) + 2\mu y_0 + 2)x_1(t)}{(1+\mu(x_1(t)+y_0))^2} \tag{89}$$

$$T_2(x) = -\frac{hL\mu}{2m(1+\mu(x_1(t)+y_0))^2}. \tag{90}$$

With the given parameters and the safety set C_s, the scheduling variables $T_1(x)$ and $T_2(x)$ meet the conditions $-0.0206 \le T_1(x) \le 0.0157$ and $-0.0676 \le T_2(x) \le -0.0470$, respectively. Due to the extreme values of $T_1(x)$ and $T_2(x)$, the LPV representation of the nonlinear system (87) requires four vertices, which corresponds to $2^p = 4$. The unknown values of $A_{\mathbf{i}}$ and $B_{\mathbf{i}}$, for all $\mathbf{i} \in \mathbb{B}^2$, are

$$A_{00} = A_{01} = \begin{bmatrix} 1.0000 & 0.0100 \\ -0.0206 & 0.9990 \end{bmatrix} \tag{91}$$

$$A_{10} = A_{11} = \begin{bmatrix} 1.0000 & 0.0100 \\ 0.0157 & 0.9990 \end{bmatrix} \tag{92}$$

$$B_{00} = B_{10} = \begin{bmatrix} 0 \\ -0.0676 \end{bmatrix} \tag{93}$$

$$B_{01} = B_{11} = \begin{bmatrix} 0 \\ -0.0470 \end{bmatrix}. \tag{94}$$

Furthermore, the gain-scheduled variables are defined as $\omega_0^1 = \frac{0.0157 - T_1(x)}{0.0363}$, $\omega_1^1 = 1 - \omega_0^1$, $\omega_0^2 = \frac{-0.0470 - T_2(x)}{0.0206}$, and $\omega_1^2 = 1 - \omega_0^2$. It is worth noting that $T_1(x)$ can be expressed as a function of ω_0^1 and ω_1^1 using the equation $T_1(x) = -0.0206\omega_0^1 + 0.0157\omega_1^1$. Similarly, one gets $T_2(x) = -0.0676\omega_0^2 - 0.0470\omega_1^2$. In order to conduct the simulation, it is presumed that the noise $w(t)$ follows a Gaussian distribution and has a covariance of $0.00001I$, with $\lambda = 0.90$ and $\delta = 0.10$.

To ensure a fair comparison and highlight the robustness of the presented minimum-variance method, it is important to note that the safe control approach outlined in Theorem 2 is performed without the inclusion of any noise measurements. This particular approach is referred to as the certainty-equivalence safe control method. By adopting this approach, one can assess the core performance of the minimum-variance method and its capability to handle challenging conditions, such as noisy environments, without any additional considerations or adjustments.

It should be pointed out that as stated in Remark 2, indirect learning methods like [28] require $2N_v(n+m) = 2\times4\times(2+1) = 24$ independent samples to design a control policy for the system (87), while the presented data-based approach in Theorem 3, which bypasses system identification, only requires $2N_v n = 2\times4\times2 = 16$ independent samples to learn a safe control policy.

Fig. 2 illustrates the state trajectories of the closed-loop system, starting from $x = [-0.0333, 0.5071]^T$, and computed for 100 distinct realizations of the noise for both safe control learning methods. By analyzing this figure, it is evident that utilizing the designed minimum variance-based approach

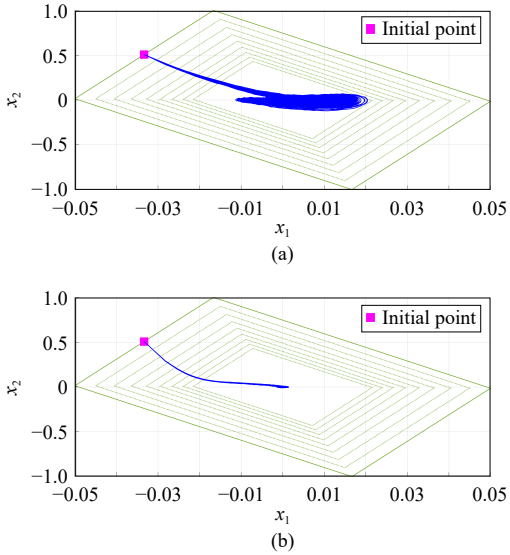

Fig. 2. Time history of the system trajectory for 100 realizations of the Gaussian noise with $\Sigma = 0.00001I$ using a) certainty-equivalence safe control and b) minimum variance-based probabilistic safe control.

results in system trajectories that remain within the bounds of the safe set with a probability of at least $1-\delta$, and do not exceed them. In contrast, the system trajectories using the certainty-equivalence method have the potential to violate the safety conditions. This indicates that the presented safe controller is more robust in the presence of noise.

Fig. 3 displays the progression of the system's states for one of the realizations of $w(t)$ utilizing the certainty-equivalence and the presented probabilistic safe controllers. It should be pointed out that no control approach can guarantee exact convergence to the equilibrium point in the presence of noise. The presented approach offers probabilistic convergence, which provides more predictability and reduced variance around the equilibrium point, as illustrated in Fig. 3. Additionally, the simulation results demonstrate that after one second, the system states converge to a vicinity of the origin with high probability. Also, time traces of the system input under both methods are exhibited in Fig. 4.

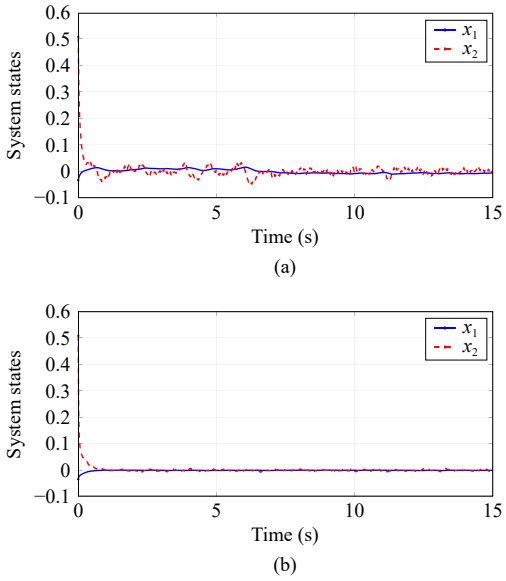

Fig. 3. Time history of the system states for 100 realizations of the Gaussian noise with $\Sigma = 0.00001I$ using a) certainty-equivalence safe control and b) minimum variance-based probabilistic safe control.

To further demonstrate the effectiveness of the established probabilistic safe controller, another Gaussian noise with a covariance of $0.0001I$ is applied to the system described in (87). The initial values of the system states are set to $x = [0.0167, -1]^T$. The phase portrait of the system states for 100 different realizations of the noise are exhibited in Fig. 5. The result clearly indicates that the presented approach exhibits significantly higher robustness in noisy environments compared to the certainty-equivalence method, which is prone to safety violations. Therefore, the superiority of the probabilistic safe gain-scheduling controller is further highlighted through this simulation.

Additionally, to better verify the performance of the presented approach, Figs. 6 and 7 provide the results of applying the method established in [46] to the system (87). Specifically, Fig. 6 is compared to the results shown in Fig. 2(b) starting from the same initial condition, and Fig. 7 is compared to

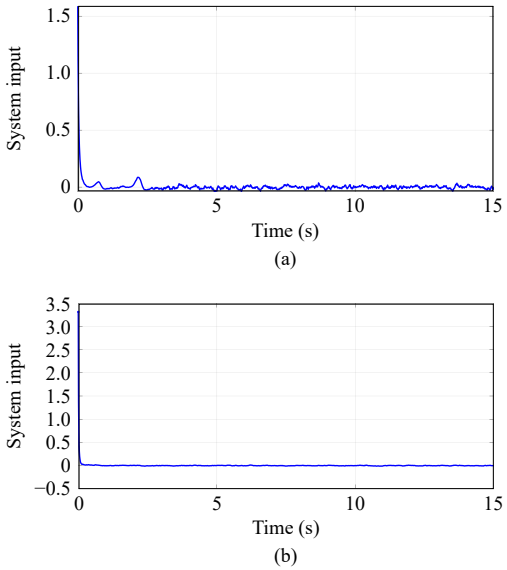

Fig. 4. Time history of the system input for 100 realizations of the Gaussian noise with $\Sigma = 0.00001I$ using a) certainty-equivalence safe control and b) minimum variance-based probabilistic safe control.

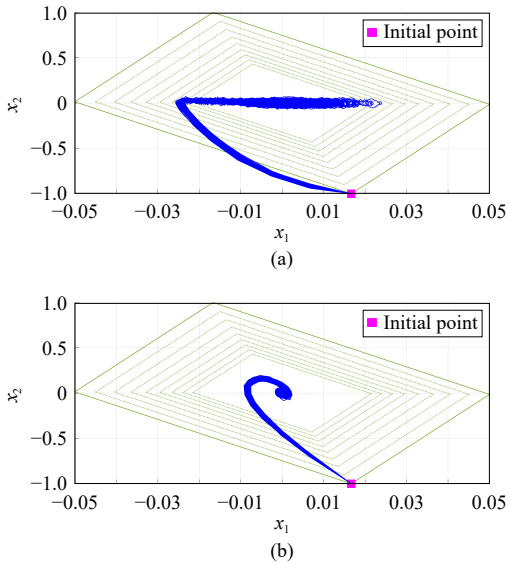

Fig. 5. Time history of the system trajectory for 100 realizations of the Gaussian noise with $\Sigma = 0.0001I$ using a) certainty-equivalence safe control and b) minimum variance-based probabilistic safe control.

those in Fig. 5(b) starting from the same initial condition. The results clearly demonstrate that the designed probabilistic safe control significantly reduces the variance of the closed-loop system in the presence of noise, thereby underscoring its superior performance.

B. Example 2: Autonomous Vehicle

Kinematic model of an autonomous vehicle is considered as follows [34]:

$$\begin{aligned} \dot{x}_e &= \omega_k y_e + v_d \cos\theta_e - v_x \\ \dot{y}_e &= -\omega_k x_e + v_d \sin\theta_e \\ \dot{\theta}_e &= \omega_{k_d} - \omega_k \end{aligned} \tag{95}$$

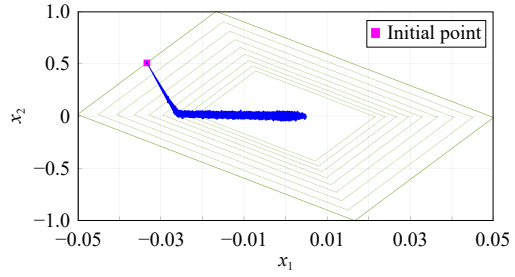

Fig. 6. Time history of the system trajectory for 100 realizations of Gaussian noise with covariance $\Sigma = 0.00001I$ using the safe control method presented in [46]. This can be compared to the result of the control method designed in this paper, as shown in Fig. 2(b), starting from the same initial condition.

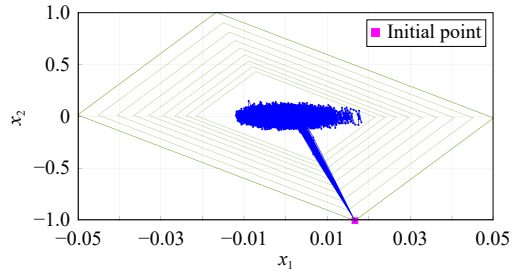

Fig. 7. Time history of the system trajectory for 100 realizations of Gaussian noise with covariance $\Sigma = 0.0001I$ using the safe control method presented in [46]. This can be compared to the result of the control method designed in this paper, as shown in Fig. 5(b), starting from the same initial condition.

in which x_e, y_e, and θ_e denote the errors in position and orientation. The variables v_x and ω_k represent the longitudinal and angular velocities, while v_d and ω_{k_d} are the reference velocities for longitudinal and angular motion, respectively. It is important to note that the autonomous vehicle's kinematic model is open-loop unstable, requiring careful control design to ensure stability and safety. By defining T_s as the sampling time, the vector of scheduling variables as $\Gamma(t) := [\omega_k(t), v_d(t), \theta_e(t)]$, where each component is constrained within $\omega_k \in [-1.417, 1.417]$ rad/s, $v_d \in [0.1, 18]$ m/s, and $\theta_e \in [-0.0873, 0.0873]$ rad, and the state, input, and reference vectors as

$$x = \begin{bmatrix} x_e \\ y_e \\ \theta_e \end{bmatrix}, \quad u = \begin{bmatrix} v_x \\ \omega_k \end{bmatrix}, \quad r_d = \begin{bmatrix} v_d \cos\theta_e \\ \omega_{k_d} \end{bmatrix}. \tag{96}$$

The nonlinear kinematic model transforms into a discrete-time LPV representation, as expressed below:

$$x(t+1) = A(\Gamma(t))x(t) + Bu(t) - Br_d(t) \tag{97}$$

where

$$A(\Gamma(t)) = \begin{bmatrix} 1 & \omega_k T_s & 0 \\ -\omega_k T_s & 1 & v_d \dfrac{\sin(\theta_e)}{\theta_e} T_s \\ 0 & 0 & 1 \end{bmatrix}, \; B = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 0 & -1 \end{bmatrix} T_s. \tag{98}$$

According to Remark 1, the system (97) can be expressed as the following LPV model:

$$x(t+1) = \sum_{r=1}^{N_v} \omega_i (A_i x(t) + B(u(t) - r_d(t))). \tag{99}$$

Defining $\bar{u}(t) := u(t) - r_d(t) = \sum_{i=1}^{n_v} K_i x(t)$, one gets

$$x(t+1) = \sum_{r=1}^{N_v} \omega_i (A_i + BK_i) x(t). \tag{100}$$

Here, number of vertices, i.e., N_v is computed as $N_v = 2^{N_g}$, where N_g shows the number of gain-scheduling variables. Furthermore, the gain-scheduling weights are computed as

$$\omega_r = \prod_{s=1}^{N_g} \nu_{rs}(\omega_0^s, \omega_1^s), \; r = 1, \ldots, N_v \tag{101}$$

with

$$\omega_0^s = \frac{\Gamma_s^u - \Gamma_s(t)}{\Gamma_s^u - \Gamma_s^l} \tag{102}$$

$$\omega_1^s = 1 - \omega_0^s, \; s = 1, \ldots, N_g \tag{103}$$

where $\Gamma_s(t)$ is the s-th component of the vector of gain-scheduling variables, i.e., $\Gamma(t)$, and Γ_s^l and Γ_s^u denote their corresponding lower and upper bounds, respectively. Also, $\nu_{rs}(\omega_0^s, \omega_1^s)$ signifies any weighting function associated with each rule r for $r = 1, \ldots, N_v$.

Since, in this example, there are three gain scheduling variables, i.e., $\Gamma(t) = [\omega_k(t), v_d(t), \theta_e(t)]$, and each of them are bounded with their corresponding lower and upper bounds, one has $N_v = 2^3 = 8$. Each of the polytopic vertex systems, i.e., A_i, is derived as a combination of the extreme values of the scheduling variables [47].

The admissible set for the system states is specified as $C_s = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid |x_1| \le 0.2, |x_2| \le 0.5, |x_3| \le 0.0873\}$.

It's important to note that, as mentioned in Remark 2, indirect learning techniques such as those discussed in [28] necessitate the collection of $2N_v(n+m) = 2 \times (3+2) = 80$ independent samples to create a control strategy for the system defined in equation (95). In contrast, the data-driven approach introduced in Theorem 3, which avoids the need for system identification, only mandates the gathering of $2N_v n = 2 \times 8 \times 3 = 48$ independent samples to acquire the information needed for developing a safe control policy.

During the simulation phase, all states are subject to Gaussian noise with a variance of 0.01. We then apply the acquired data-based probabilistic safe control policy to the system described in (95). Fig. 8 provides a time-based representation of the system states, while Fig. 9 displays the trajectory of the autonomous vehicle operating in a noisy environment. Notably, these visualizations reveal that the autonomous vehicle successfully achieves safe path tracking even in the presence of noise, without encountering any safety violations such as collisions with road boundaries[1].

[1] To watch the animation of the path tracking performance, please click on the following link: Safe path tracking animation.

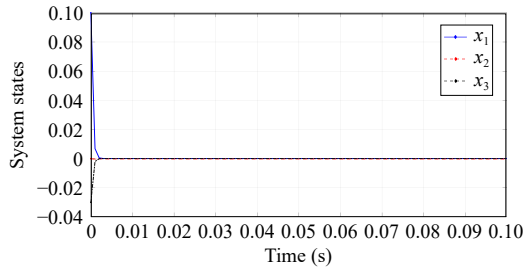

Fig. 8. Time evolution of the states for the system (95).

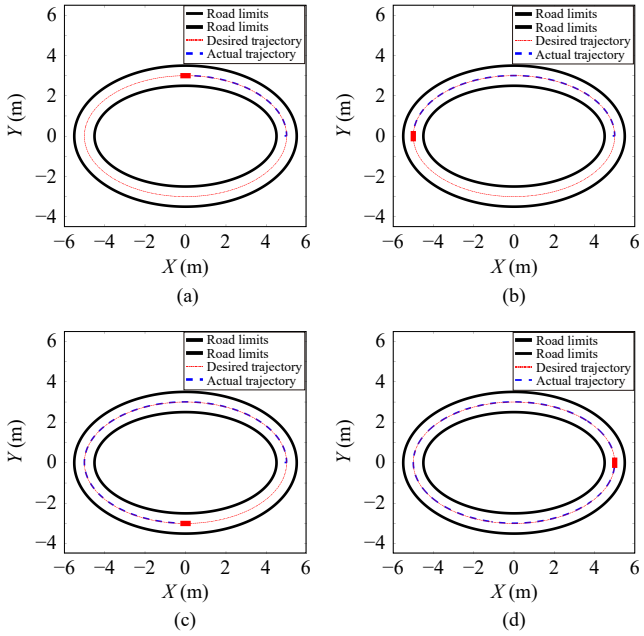

Fig. 9. Time evolution of the vehicle's trajectory (the red rectangle represents the vehicle).

C. Experimental Implementation: Autonomous Vehicle

In this part, the designed controller is implemented on an actual autonomous vehicle, named the ROSbot 2R robot, as shown in Fig. 10. The robot operating system (ROS) is used to control its kinematics. The robot environment is illustrated in Fig. 11, where the black boxes represent obstacles and the polyhedral safe set is defined as the 2D space between them. The black dot on the floor indicates the desired set-point for the robot. The kinematic model of this robot is as follows:

$$\begin{aligned}\dot{x} &= v\cos\theta\\ \dot{y} &= v\sin\theta\\ \dot{\theta} &= \omega_b.\end{aligned} \tag{104}$$

In this model, x and y represent the robot's positions, and θ denotes its orientation. Additionally, v and ω_b are the linear and angular velocities of the robot, respectively. By defining $x_s = [x, y, \theta]^T$ as the state vector and $x_{s,d} = [x_d, y_d, \theta_d]^T$ as the desired state, the discrete-time error model of the kinematic system is obtained as follows:

$$x_e(t+1) = Ax_e(t) + B(\theta_e + \theta_d)u_s(t) \tag{105}$$

where

Fig. 10. The ROSbot 2R robot.

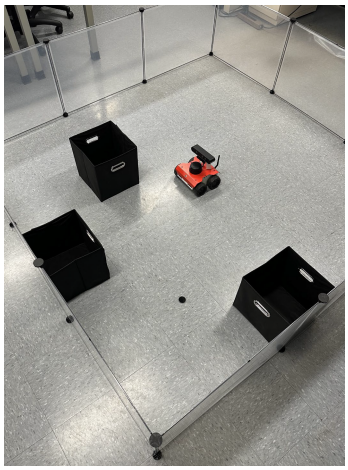

Fig. 11. Robot's environment.

$$A = I,\ B(\theta_e+\theta_d) = \begin{bmatrix}\cos(\theta_e+\theta_d) & 0\\ \sin(\theta_e+\theta_d) & 0\\ 0 & 1\end{bmatrix} T_s \tag{106}$$

and $x_e = x_s - x_{s,d} = [x_e, y_e, \theta_e]^T$ is the error system's state. The objective is to stabilize the error system, which is equivalent to tracking a desired set-point.

Initially, the environment and the controller are implemented in Gazebo. The dynamic behavior of the robot in Gazebo is illustrated in Fig. 12 with different episodes of the simulation over time[2]. Subsequently, a real implementation on the robot, as shown in Fig. 10, is conducted. Different episodes of this real implementation are presented in Fig. 13 to provide a visual representation of the robot's performance over time[3]. It is demonstrated that, under the designed probabilistic safe control approach, the robot can effectively follow the desired set-point in practice. In the experimental validation, no significant computational burden is encountered due to the fact that a linear programming optimization problem is

[2] To watch the video of the set-point tracking performance of the robot in Gazebo, please click on the following link: Safe set-point tracking simulation.

[3] To watch the video of the set-point tracking performance of the robot in real-world, please click on the following link: Safe set-point tracking implementation.

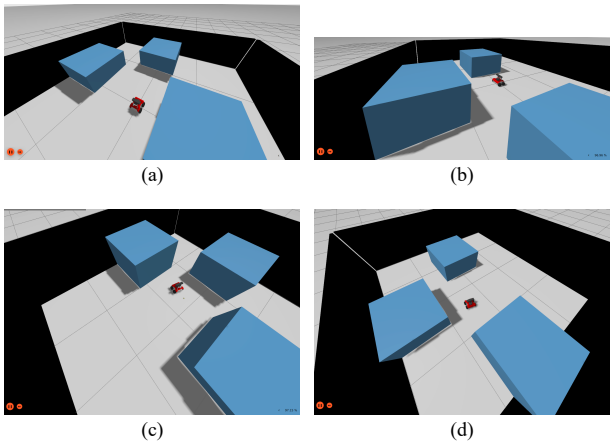

Fig. 12. Time evolution of the vehicle's motion in the Gazebo simulation.

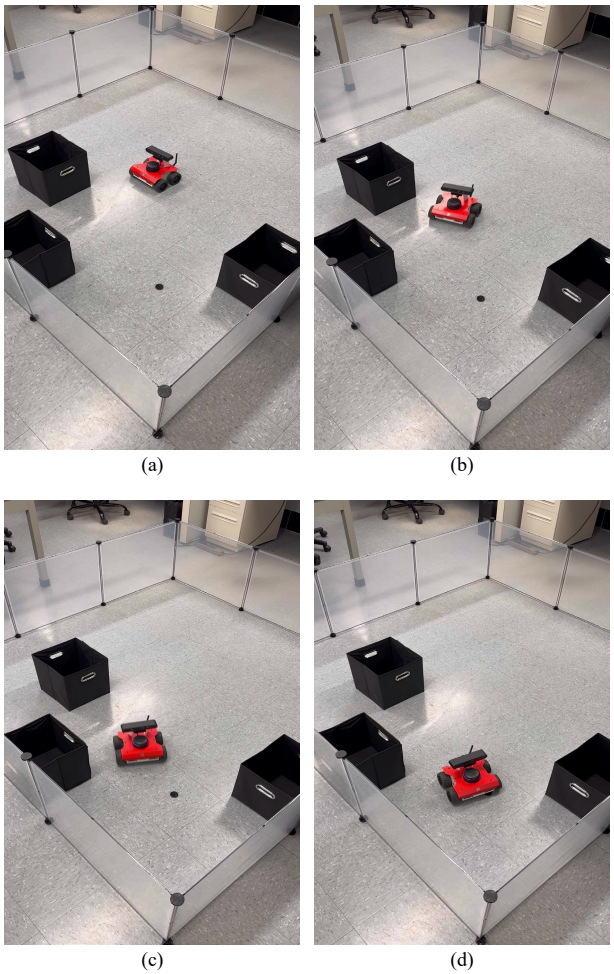

Fig. 13. Time evolution of the vehicle's motion in real-world implementation (the black boxes are obstacles, and the black dot on the floor is the desired-point).

solved once using the data collected offline to calculate the LPV gains. These gains are then employed in the implementation, resulting in a considerable decrease in computational requirements, particularly in practical scenarios.

VI. Conclusion

This paper presents a risk-informed model-free safe controller for nonlinear discrete-time systems. A probabilistic safe gain-scheduling controller is directly learned from data to bypass the system identification and learn a safe controller while only relying on measured data. The presented data-based safe control design amounts to a numerically-efficient linear program for polyhedral safe sets. In comparison with the existing indirect approaches, a minimum-variance direct safe controller that takes into account the risk of safety violation is learned using less data, and provides robustness guarantees.

A limitation of the presented approach, in contrast to the existing barrier certifier approach, is that it is limited to only convex safe sets. Future work will extend these results to systems with nonconvex safe sets. Furthermore, the designed safe control scheme will be integrated with reinforcement learning-based controllers to certify their safety while minimizing interference with their actions. Additionally, for situations where the LPV approximation is inadequate and significant nonlinear dynamics persist, our future work will aim to develop a combined LPV and explicit nonlinear representation approach. This will provide a more accurate depiction of the system's dynamics. We will then design a direct data-driven safety control strategy that utilizes the LPV framework while also addressing residual nonlinearities in the closed-loop behavior.

References

[1] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annu. Rev. Control Robot. Auton. Syst.*, vol. 5, pp. 411–444, May 2022.

[2] M. Zanon and S. Gros, "Safe reinforcement learning using robust MPC," *IEEE Trans. Autom. Control*, vol. 66, no. 8, pp. 3638–3652, Aug. 2021.

[3] R. Grandia, A. J. Taylor, A. D. Ames, and M. Hutter, "Multi-layered safety for legged robots via control barrier functions and model predictive control," in *Proc. IEEE Int. Conf. Robotics and Automation*, Xi'an, China, 2021, pp. 8352-8358.

[4] M. Mazouchi, S. Nageshrao, and H. Modares, "Conflict-aware safe reinforcement learning: A meta-cognitive learning framework," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 3, pp. 466–481, Mar. 2022.

[5] L. Zhang, R. Zhang, T. Wu, R. Weng, M. Han, and Y. Zhao, "Safe reinforcement learning with stability guarantee for motion planning of autonomous vehicles," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5435–5444, Dec. 2021.

[6] M. Alshiekh, R. Bloem, R. Ehlers, B. Könighofer, S. Niekum, and U. Topcu, "Safe reinforcement learning via shielding," in *Proc. 32nd AAAI Conf. Artificial Intelligence*, New Orleans, USA, 2018, pp. 2669–2678.

[7] S. Li and O. Bastani, "Robust model predictive shielding for safe reinforcement learning with stochastic dynamics," in *Proc. IEEE Int. Conf. Robotics and Automation*, Paris, France, 2020, pp. 7166–7172.

[8] S. Gao, Z. Peng, H. Wang, L. Liu, and D. Wang, "Safety-critical model-free control for multi-target tracking of USVs with collision avoidance," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 7, pp. 1323–1326, Jul. 2022.

[9] G. Yang, C. Belta, and R. Tron, "Self-triggered control for safety critical systems using control barrier functions," in *Proc. American Control Conf.*, Philadelphia, USA, 2019, pp. 4454–4459.

[10] Z. Marvi and B. Kiumarsi, "Barrier-certified learning-enabled safe control design for systems operating in uncertain environments," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 3, pp. 437–449, Mar. 2021.

[11] M. Ahmadi, X. Xiong, and A. D. Ames, "Risk-averse control via CVaR barrier functions: Application to bipedal robot locomotion," *IEEE Control Syst. Lett.*, vol. 6, pp. 878–883, 2022.

[12] R. Cheng, G. Orosz, R. M. Murray, and J. W. Burdick, "End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks," in *Proc. 33rd AAAI Conf. Artificial*

Intelligence, Honolulu, USA, 2019, pp. 3387–3395.

[13] J. Zeng, B. Zhang, and K. Sreenath, "Safety-critical model predictive control with discrete-time control barrier function," in *Proc. American Control Conf.*, New Orleans, USA, 2021, pp. 3882–3889.

[14] J. Seo, J. Lee, E. Baek, R. Horowitz, and J. Choi, "Safety-critical control with nonaffine control inputs via a relaxed control barrier function for an autonomous vehicle," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 1944–1951, Apr. 2022.

[15] A. Chern, X. Wang, A. Iyer, and Y. Nakahira, "Safe control in the presence of stochastic uncertainties," in *Proc. 60th IEEE Conf. Decision and Control*, Austin, USA, 2021, pp. 6640–6645.

[16] A. Agrawal and K. Sreenath, "Discrete control barrier functions for safety-critical control of discrete systems with application to bipedal robot navigation," in *Proc. Robotics: Science and Systems*, Cambridge, USA, 2017.

[17] S. Samuelson and I. Yang, "Safety-aware optimal control of stochastic systems using conditional value-at-risk," in *Proc. Annu. American Control Conf.*, Milwaukee, USA, 2018, pp. 6285–6290.

[18] M. F. Reis, A. P. Aguiar, and P. Tabuada, "Control barrier function-based quadratic programs introduce undesirable asymptotically stable equilibria," *IEEE Control Syst. Lett.*, vol. 5, no. 2, pp. 731–736, Apr. 2021.

[19] Y. Lan and I. Mezić, "Linearization in the large of nonlinear systems and Koopman operator spectrum," *Phys. D Nonlinear Phenom.*, vol. 242, no. 1, pp. 42–53, Jan. 2013.

[20] B. Esmaeili, M. Salim, and M. Baradarannia, "Predefined performance-based model-free adaptive fractional-order fast terminal sliding-mode control of MIMO nonlinear systems," *ISA Trans.*, vol. 131, pp. 108–123, Dec. 2022.

[21] B. Esmaeili, S. S. Madani, M. Salim, M. Baradarannia, and S. Khanmohammadi, "Model-free adaptive iterative learning integral terminal sliding mode control of exoskeleton robots," *J. Vib. Control*, vol. 28, no. 21-22, pp. 3120–3139, Nov. 2022.

[22] F. Blanchini and S. Miani, *Set-Theoretic Methods in Control*. Boston, USA: Springer, 2008.

[23] Z. Gao and J. Fu, "Robust LPV modeling and control of aircraft flying through wind disturbance," *Chin. J. Aeronaut.*, vol. 32, no. 7, pp. 1588–1602, Jul. 2019.

[24] K. Zhu, D. Ma, and J. Zhao, "Event triggered control for a switched LPV system with applications to aircraft engines," *IET Control Theory Appl.*, vol. 12, no. 10, pp. 1505–1514, Jul. 2018.

[25] A. San-Miguel, V. Puig, and G. Alenyà, "Disturbance observer-based LPV feedback control of a *N*-DoF robotic manipulator including compliance through gain shifting," *Control Eng. Pract.*, vol. 115, p. 104887, Oct. 2021.

[26] P. S. G. Cisneros, A. Sridharan, and H. Werner, "Constrained predictive control of a robotic manipulator using quasi-LPV representations," *IFAC-PapersOnLine*, vol. 51, no. 26, pp. 118–123, 2018.

[27] M. A. H. Darwish, P. B. Cox, I. Proimadis, G. Pillonetto, and R. Tóth, "Prediction-error identification of LPV systems: A nonparametric Gaussian regression approach," *Automatica*, vol. 97, pp. 92–103, Nov. 2018.

[28] A. Bisoffi, C. De Persis, and P. Tesi, "Data-based guarantees of set invariance properties," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 3953–3958, 2020.

[29] A. Luppi, C. De Persis, and P. Tesi, "On data-driven stabilization of systems with nonlinearities satisfying quadratic constraints," *Syst. Control Lett.*, vol. 163, p. 105206, May 2022.

[30] A. Bisoffi, C. De Persis, and P. Tesi, "Controller design for robust invariance from noisy data," *IEEE Trans. Autom. Control*, vol. 68, no. 1, pp. 636–643, Jan. 2023.

[31] C. De Persis and P. Tesi, "Low-complexity learning of linear quadratic regulators from noisy data," *Automatica*, vol. 128, p. 109548, Jun. 2021.

[32] H. Modares, "Data-driven safe control of uncertain linear systems under aleatory uncertainty," *IEEE Trans. Autom. Control*, vol. 69, no. 1, pp. 551–558, Jan. 2024.

[33] J. M. Snider, "Automatic steering methods for autonomous automobile path tracking," Carnegie Mellon University, Pittsburgh, USA, Tech. Rep. CMU-RI-TR-09-08, 2009.

[34] E. Alcala, V. Puig, J. Quevedo, and T. Escobet, "Gain-scheduling LPV control for autonomous vehicles including friction force estimation and compensation mechanism," *IET Control Theory Appl.*, vol. 12, no. 12, pp. 1683–1693, Aug. 2018.

[35] E. Alcalá, V. Puig, and J. Quevedo, "LPV-MPC control for autonomous vehicles," *IFAC-PapersOnLine*, vol. 52, no. 28, pp. 106–113, 2019.

[36] X. Geng and L. Xie, "Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization," *Annu. Rev. Control*, vol. 47, pp. 341–363, 2019.

[37] A. Marcos and G. J. Balas, "Development of linear-parameter-varying models for aircraft," *J. Guid. Control Dyn.*, vol. 27, no. 2, pp. 218–228, Mar. 2004.

[38] P. H. S. Coutinho, M. L. C. Peixoto, I. Bessa, and R. M. Palhares, "Dynamic event-triggered gain-scheduling control of discrete-time quasi-LPV systems," *Automatica*, vol. 141, p. 110292, Jul. 2022.

[39] H. S. Abbas, R. Tóth, M. Petreczky, N. Meskin, and J. Mohammadpour, "Embedding of nonlinear systems in a linear parameter-varying representation," *IFAC Proc. Vol.*, vol. 47, no. 3, pp. 6907–6913, 2014.

[40] E. Kofman, J. A. De Doná, and M. M. Seron, "Probabilistic set invariance and ultimate boundedness," *Automatica*, vol. 48, no. 10, pp. 2670–2676, Oct. 2012.

[41] S. Mate, H. Kodamana, S. Bhartiya, and P. S. V. Nataraj, "A stabilizing sub-optimal model predictive control for quasi-linear parameter varying systems," *IEEE Control Syst. Lett.*, vol. 4, no. 2, pp. 402–407, Apr. 2020.

[42] P. Coppens, M. Schuurmans, and P. Patrinos, "Data-driven distributionally robust LQR with multiplicative noise," in *Proc. 2nd Conf. Learning for Dynamics and Control*, Berkeley, USA, 2020, pp. 521–530.

[43] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 2012.

[44] T. Lattimore and C. Szepesvári, *Bandit Algorithms*. Cambridge, UK: Cambridge University Press, 2020.

[45] M. L. C. Peixoto, M. F. Braga, and R. M. Palhares, "Gain-scheduled control for discrete-time non-linear parameter-varying systems with time-varying delays," *IET Control Theory Appl.*, vol. 14, no. 19, pp. 3217–3229, Dec. 2020.

[46] A. Modares, N. Sadati, B. Esmaeili, F. A. Yaghmaie, and H. Modares, "Safe reinforcement learning via a model-free safety certifier," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 3302–3311, Mar. 2024.

[47] A. Kwiatkowski, M.-T. Boll, and H. Werner, "Automated generation and assessment of affine LPV models," in *Proc. 45th IEEE Conf. Decision and Control*, San Diego, USA, 2006, pp. 6690–6695.

Babak Esmaeili (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering and control and the M.Sc. degree in mechatronics engineering from the University of Tabriz, Iran, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Department of Mechanical Engineering, Michigan State University, USA. His current research interests include data-based control, safe reinforcement learning, and autonomous systems. Mr. Esmaeili serves as a Reviewer for several international journals, including *IEEE Transactions on Cybernetics*, *ISA Transactions*, and *Journal of Vibration and Control*.

Hamidreza Modares (Senior Member, IEEE) received the B.S. degree from the University of Tehran, Iran, in 2004, the M.S. degree from the Shahrood University of Technology, Iran, in 2006, and the Ph.D. degree from The University of Texas at Arlington, USA, in 2015. He was a Senior Lecturer with the Shahrood University of Technology from 2006 to 2009 and a Faculty Research Associate with The University of Texas at Arlington from 2015 to 2016. He is currently an Assistant Professor with the Department of Mechanical Engineering, Michigan State University, USA. His current research interests include cyber-physical systems, reinforcement learning, distributed control, robotics, and machine learning. Dr. Modares was a recipient of the Best Paper Award from the 2015 IEEE International Symposium on Resilient Control Systems. He is also an Associate Editor of the *IEEE Transactions on Neural Networks and Learning Systems*, *IEEE Transactions on Systems, Man, and Cybernetics* (SMC), and *Neurocomputing*.