

Data-Driven Safe Control of Stochastic Nonlinear Systems

Babak Esmaili* Hamidreza Modares**

* *Mechanical Engineering Department, Michigan State University, East Lansing, MI 48824, USA (e-mail: esmaeil1@msu.edu).*

** *Mechanical Engineering Department, Michigan State University, East Lansing, MI 48824, USA (e-mail: modares@msu.edu)*

Abstract: This paper introduces a data-based safe control design for stochastic nonlinear systems. The controller consists of two parts: a linear component that ensures set invariance, and a nonlinear component that minimizes the impact of nonlinearities. A closed-form presentation of both dynamics is provided, linking control gains directly to data and decision variables. The design is formulated as a semidefinite programming problem (SDP) to ensure robust set invariance, incorporating extra constraints for unmeasured noise and residual nonlinear effects. The effectiveness of this approach is validated through simulations, demonstrating its potential to enhance safety and performance in controlling nonlinear dynamics.

Copyright © 2024 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: Data-Driven Control, Safety, Stochastic, Nonlinearity Minimization

1. INTRODUCTION

The success of learning-enabled autonomous systems has the potential to bring unprecedented benefits to humanity and can make a profound impact on variety of engineering sectors, including unmanned vehicles and robots. Despite the rapid progress on learning-enabled systems, tension is also increasing in the community from a lack of safety guarantees. Ensuring the safety of autonomous stochastic nonlinear systems remains a formidable challenge, as traditional methods often resort to simplified linear models, which may not adequately capture the intricate dynamics of these systems (De Persis et al., 2023; Guo et al., 2023). Although recent advancements in data-driven control suggest potential solutions (Hou and Wang, 2013), fully integrating robust safety measures into control designs for systems with unknown nonlinear dynamics and limited data availability is still challenging.

Safety certifications are particularly crucial for reinforcement learning (RL), as key to learning-enabled autonomy. Existing safety certificates typically rely on control barrier functions (CBFs), which provide safety guarantees but may lead to sub-optimal performance due to their myopic nature (Reis et al., 2020; Brunke et al., 2022; Zanon and Gros, 2020; Grandia et al., 2021; Mazouchi et al., 2021; Zhang et al., 2021; Alshiekh et al., 2018; Li and Bastani, 2020). This highlights the need for a model-free approach where safe and RL controllers are directly learned from data, circumventing the limitations inherent in CBF methods that depend on an explicit system model for the safety certificate.

To address nonlinear systems, several approaches have been explored, such as dynamic linearization (Esmaili et al., 2022) or modeling them as linear parameter varying (LPV) systems (Lan and Mezić, 2013). Despite these methodologies, research on

designing safe controllers for nonlinear discrete-time systems remains sparse. Some efforts have focused on leveraging the LPV framework to develop efficient safe controllers for a specific class of nonlinear systems, as demonstrated in (Modares et al., 2023). These methods, however, limit the class of nonlinear systems and can be data hungry. Besides, they often do not fully address aleatory uncertainty, treating it as a bounded disturbance and providing robust but conservative safety assurances (Ahmadi et al., 2020; Bisoffi et al., 2020). Modeling the noise with a bounded disturbance is particularly insufficient in systems where noise exhibits infinite support, highlighting the urgent need for robust data-based safe controllers that can effectively manage variability in safety violations and reduce performance fluctuations.

The paper introduces a novel data-based control design methodology tailored for autonomous stochastic nonlinear systems, emphasizing safety and robustness. This method enhances the conventional data-driven framework by incorporating stochastic elements and noise, which are typically absent in standard models such as those discussed in (De Persis et al., 2023). Unlike prior models that do not account for stochastic behavior and noise, this approach explicitly integrates these factors into the control process, thereby improving safety and reliability in control applications. The controller architecture combines a linear component designed to maintain invariant safety sets through λ -contractivity, with a nonlinear component crafted to mitigate or neutralize the effects of nonlinearities on set invariance. This comprehensive strategy is articulated through a closed-loop data-based representation of both linear and nonlinear dynamics. Here, the corresponding control gains are defined directly in terms of data and decision variables, streamlining the design process. The formulation of this control system is expressed as a semidefinite programming (SDP) problem. This framework allows for the learning of decision variables that ensure robust set invariance, addressing the challenges posed by unmeasured noise and residual nonlinearities. The design

* This work is supported in part by the Department of Navy award N00014-22-1-2159 and in part by the National Science Foundation under award ECCS-2227311.

process incorporates additional constraints into the SDP to strengthen the system's resilience against these uncertainties. This method significantly reduces the data requirements traditionally necessary for system stabilization, making it more practical for real-world applications where extensive data may not be readily available. The efficacy of this approach is demonstrated through simulations, highlighting its capability to enhance both the safety and performance of stochastic nonlinear systems by adeptly managing the complexities introduced by nonlinear dynamics. The simulations reveal the method's robustness in maintaining set invariance and reducing the impact of stochastic disturbances and nonlinear interactions, affirming its practical utility and potential for real-world applications.

Notations: In this paper, I denotes the identity matrix of suitable dimension. For a matrix A , A_i indicates its i -th row, and A_{ij} represents the element located in the i -th row and j -th column. For matrices or vectors A and B of identical dimensions, $A(\leq, \geq)B$ signifies a component-wise comparison, meaning $A_{ij}(\leq, \geq)B_{ij}$ for every i and j . $\text{Tr}(A)$ stands for the trace of matrix A , and $\|A\|_2$ and $\|A\|_F$ signify its spectral and the Frobenius norm, respectively. Also, $\log\det(A)$ denotes the logarithm of the determinant of matrix A . If Q is a matrix, $Q(\preceq, \succeq)0$ implies that Q is negative or positive semi-definite, respectively. For a set $\mathcal{S}$ and a non-negative scalar μ , $\mu\mathcal{S}$ is defined as the set containing all elements μx such that $x \in \mathcal{S}$. Additionally, it is assumed that all random variables are defined on a probability space represented as $(\Gamma, \mathcal{F}, \mathbb{P})$, where Γ is the sample space, $\mathcal{F}$ is the associated σ -algebra, and $\mathbb{P}$ is the probability measure. For a random variable $v : \Gamma \rightarrow \mathbb{R}^n$ defined on the probability space $(\Gamma, \mathcal{F}, \mathbb{P})$, the notation $v \in \mathbb{R}^n$ signifies its dimension. The mathematical expectation of v is denoted by $\mathbb{E}[v]$, and if it is known that $\mathbb{E}[v] = \hat{v}$, then the covariance of v is computed using the formula $\mathbb{E}[(v - \hat{v})(v - \hat{v})^T]$.

Lemma 1. ((Coppens et al., 2020)). For a specified random vector $v \in \mathbb{R}^{n \times 1}$ and a matrix $Q \in \mathbb{R}^{n \times n}$, one has

$$\mathbb{E}[v^T Q v] = \text{Tr}(Q \mathbb{E}[\tilde{v} \tilde{v}^T]) + \mathbb{E}[v]^T Q \mathbb{E}[v], \quad (1)$$

where $\tilde{v} = v - \mathbb{E}[v]$.

Definition 1. ((Blanchini and Miani, 2008)). A C-set is characterized as a set that exhibits both convexity and compactness, and its interior includes the origin.

Definition 2. ((Blanchini and Miani, 2008)). A polyhedral C-set, represented by $\mathcal{S}(F, g)$, is defined as

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

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

Definition 3. ((Blanchini and Miani, 2008)). For a specified positive definite matrix P , an ellipsoidal C-set is represented by

$$\mathcal{E}(P, 1) = \{x \in \mathbb{R}^n : x^T P^{-1} x \leq 1\}. \quad (3)$$

2. PROBLEM FORMULATION

Consider the following discrete-time stochastic nonlinear system

$$x_{t+1} = f(x_t) + Bu_t + w_t, \quad (4)$$

where $x_t \in \mathbb{R}^n$ is the state vector at time t , $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an unknown nonlinear function mapping the state vector at time t to the state vector at time $t + 1$, $B \in \mathbb{R}^{n \times m}$ is an unknown matrix representing the input matrix, $u_t \in \mathbb{R}^m$ is the control input vector at time t , and $w_t \in \mathbb{R}^n$ is the process noise vector

at time t , accounting for uncertainties or disturbances in the system. According to (De Persis et al., 2023), system (4) can be equivalently represented as follows

$$x_{t+1} = A' z'(x_t) + Bu_t + w_t, \quad (5)$$

where $z'(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^r$ denotes a vector-valued continuous function, and $A' \in \mathbb{R}^{n \times r}$ is a constant matrix.

Assumption 1. A continuous function $z : \mathbb{R}^n \rightarrow \mathbb{R}_z^n$ exists such that $z(x) = T z'(x)$ for a certain matrix $T \in \mathbb{R}^{r \times n_z}$.

Given Assumption 1, system (5) can be equivalently expressed as

$$x_{t+1} = Az(x_t) + Bu_t + w_t, \quad (6)$$

where $z(x_t)$ is a dictionary of functions used to capture the system dynamics. In this paper, $z(x_t)$ is chosen such that it contains both linear and nonlinear functions, i.e., $z(x_t) = [x_t^T S(x_t)^T]^T$, with $S(x_t)$ comprising solely the nonlinear functions. Additionally, the matrix A is partitioned as $A = [A_l A_{nl}]$, where $A_l \in \mathbb{R}^{n \times n}$ represents the matrix for linear component of $z(x_t)$, and $A_{nl} \in \mathbb{R}^{n \times (n_z - n)}$ represents its nonlinear component.

Assumption 1 means that we choose a library of functions capable of describing the dynamics of the system, a necessity in many practical cases such as mechanical and electrical systems where the dynamics can be derived from first principles, but exact system parameters may be unknown. This assumption ensures that the chosen functions $z(x_t)$ can adequately capture the system's dynamics, facilitating control design.

Assumption 2. The vector $w_t = [w_{1,t}, \dots, w_{n,t}]^T$, which represents the noise in system (4), is assumed to follow a Gaussian distribution. This means that it has a zero mean and a variance of Σ , expressed as $w \sim \mathcal{N}(0, \Sigma)$, where $\mathbb{E}[w_{i,t} w_{j,t}] = 0$ for $i \neq j$, and $\mathbb{E}[w_{i,t}^2] = \sigma_i^2$ for $i = 1, \dots, n$.

Assumption 3. The unknown matrix pair (A_l, B) is stabilizable.

Before delving into the problem formulation, the concept of contractive sets, as a fundamental tool for ensuring safety in this paper, is introduced.

Definition 4. (Contractive Set for Deterministic Systems, i.e., the system (4) with $w_t \equiv 0$): If for every $x_t \in \mathcal{S} \subseteq \mathbb{R}^n$, it holds that $x_{t+1} \in \lambda \mathcal{S}$ for all $t \geq 0$, where $0 < \lambda \leq 1$, then $\mathcal{S}$ is called a λ -contractive set.

Definition 5. ((Modares, 2023)). (Contractive Set in Expectation (CSiE)): A set $\mathcal{S} \subseteq \mathbb{R}^n$ is considered λ -contractive in expectation for the system (4) if $x_t \in \mathcal{S}$ implies that $\mathbb{E}[x_{t+1}] \in \lambda \mathcal{S}$ for all $t \geq 0$.

Definition 6. (Admissible Set): An admissible set is characterized by the permissible physical boundaries within which the system is allowed to operate.

Definition 7. (Safe Set): A subset of an admissible set is termed a safe set if, under certain conditions, the system's trajectories remain within it, ensuring safety.

Assumption 4. The permissible operational boundaries are represented by a polyhedral set as described in equation (2).

In (Bisoffi et al., 2020), it is demonstrated that for a deterministic system, a λ -contractive set is also an invariant set. Specifically, if a set $\mathcal{S}$ is λ -contractive and $x_0 \in \mathcal{S}$, then it follows that $x_t \in \mathcal{S}$ for all $t \geq 0$. In the context of stochastic systems, (Modares, 2023) establishes that if the set $\mathcal{S}$ satisfies the conditions of being CSiE, then such a set is deemed safe in expectation. Consequently, if $x_0 \in \mathcal{S}$, then $\mathbb{E}[x_t] \in \mathcal{S}$ for all $t \geq 0$.

Although system trajectories are allowed to evolve within the admissible set, it is not always feasible to ensure that the entire admissible set is safe. The size of the safe set is dependent on the quality of the data and the design of the control system. Consequently, to enhance safety, the controller is designed to maximize the size of the safe set, utilizing the available data and control structure. This approach leads to the identification of the largest invariant ellipsoid contained within the admissible set by implementing a model-free safe nonlinear state-feedback controller.

Furthermore, since the exact model of the systems' dynamics are not fully known, and to reduce the impact of modeling errors, a data-based approach will be used to design the safe controller.

The following problem summarizes the core challenge addressed by this research.

Problem 1. Let $\mathcal{S}$ be a polyhedral admissible set. The objective is to find a nonlinear state-feedback control in the form of

$$u_t = Kz(x_t) \quad (7)$$

with K as the feedback gain matrix to make the largest ellipsoid contained in the polyhedral admissible set $\mathcal{S}$ invariant.

This data-driven method allows for the incorporation of real-time data and learning techniques to adapt to the actual behavior of the system, ensuring safety even in the presence of noises.

3. CONTROL DESIGN

To construct the data-driven safe nonlinear control, we initiate the process by applying a sequence of inputs

$$U_0 := [u_0 \ u_1 \ \cdots \ u_{N-1}] \in \mathbb{R}^{m \times N} \quad (8)$$

to the system (4) and gathering $N + 1$ samples of the state vectors as follows

$$X := [x_0 \ x_1 \ \cdots \ x_N] \in \mathbb{R}^{n \times (N+1)} \quad (9)$$

The collected samples are subsequently arranged in the following manner:

$$X_0 := [x_0 \ x_1 \ \cdots \ x_{N-1}] \in \mathbb{R}^{n \times N} \quad (10)$$

$$X_1 := [x_1 \ x_2 \ \cdots \ x_N] \in \mathbb{R}^{n \times N} \quad (11)$$

$$Z_0 := \begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \\ S(x_0) & S(x_1) & \cdots & S(x_{N-1}) \end{bmatrix} \in \mathbb{R}^{n_z \times N} \quad (12)$$

Moreover, the sequence of unknown disturbances can be represented as

$$W_0 = [w_0 \ w_1 \ \cdots \ w_{N-1}] \in \mathbb{R}^{n \times N}. \quad (13)$$

The initial phase is to construct a representation of the closed-loop system based on data. Influenced by the work in (De Persis et al., 2023), and leveraging the data obtained along with the nonlinear system (4), the data-based form is presented as follows

Lemma 2. Consider the existence of matrices $K \in \mathbb{R}^{m \times n_z}$ and $G \in \mathbb{R}^{N \times n_z}$ such that

$$\begin{bmatrix} K \\ I_{n_z} \end{bmatrix} = \begin{bmatrix} U_0 \\ Z_0 \end{bmatrix} G, \quad (14)$$

where $G = [G_1 \ G_2]$ with $G_1 \in \mathbb{R}^{N \times n}$ and $G_2 \in \mathbb{R}^{N \times (n_z - n)}$.

Under these conditions, the stochastic nonlinear system (4) with the controller (7) exhibits the following closed-loop dynamics:

$$x_{t+1} = (X_1 - W_0)G_1x_t + (X_1 - W_0)G_2S(x_t) + w_t. \quad (15)$$

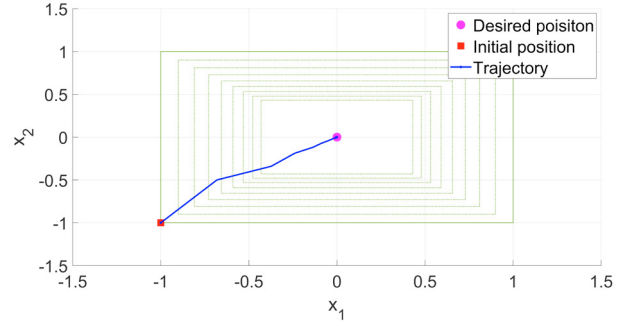


Fig. 1. Time history of the system trajectory under the designed nonlinear state-feedback controller.

Proof. By substituting $u_t = Kz(x_t)$ into (6) and employing (14), one obtains

$$\begin{aligned} x_{t+1} &= [B \ A] \begin{bmatrix} K \\ I_{n_z} \end{bmatrix} z(x_t) + w_t \\ &= [B \ A] \begin{bmatrix} U_0 \\ Z_0 \end{bmatrix} Gz(x_t) + w_t \\ &= (AZ_0 + BU_0)Gz(x_t) + w_t. \end{aligned} \quad (16)$$

Given that the data sequences U_0 , X_0 , X_1 , Z_0 , and W_0 fulfill (6), the relation $X_1 = AZ_0 + BU_0 + W_0$ is valid. By applying this to (16) and dividing G into $G = [G_1 \ G_2]$ with $G_1 \in \mathbb{R}^{N \times n}$ and $G_2 \in \mathbb{R}^{N \times (n_z - n)}$, one gets

$$\begin{aligned} x_{t+1} &= (X_1 - W_0)Gz(x_t) + w_t \\ &= (X_1 - W_0)[G_1 \ G_2] \begin{bmatrix} x_t \\ S(x_t) \end{bmatrix} + w_t. \end{aligned} \quad (17)$$

This concludes the proof. $\square$

3.1 Motivating Example

Consider the following nonlinear discrete-time system

$$x_{1,t+1} = x_{2,t} + x_{1,t}^3 + u_t, \quad (18a)$$

$$x_{2,t+1} = 0.5x_{1,t}. \quad (18b)$$

This serves as a motivating example for designing a safety-oriented nonlinear state-feedback controller. To achieve this, we choose the following dictionary of functions

$$z(x) = [x^\top \ x_1^2 \ x_2^2 \ x_1x_2 \ x_1^3 \ x_2^3 \ x_1x_2^2 \ x_1^2x_2]^\top. \quad (19)$$

where $x = [x_1 \ x_2]^\top$. The computed gain for the controller is

$$K = [0.1814 \ -0.5000 \ 0 \ -1.0000 \ 0 \ 0 \ 0 \ 0]. \quad (20)$$

With this gain, the nonlinear term $x_{1,t}^3$ in the system dynamics is effectively canceled, enhancing the safety of the control design. Figure 3 illustrates the time history of the system trajectory under the designed nonlinear state-feedback controller.

Motivated by the aforementioned example, initially, the approach assumes the availability of noise measurements, which facilitates the design of a safe nonlinear controller in Theorem 1. This initial design leverages the noise data to enhance the controller's ability to maintain system safety under stochastic disturbances. However, recognizing the restrictive nature of this assumption, the methodology is further developed to address scenarios where noise measurements are unavailable.

In Theorem 2, the strategy evolves to design a controller without relying on direct noise measurements. To mitigate the effect of ignoring the noise data and reduce the impact of noise on the closed-loop system, the revised design focuses on minimizing the variance of the closed-loop system. This optimization objective aims to enhance system robustness by diminishing the influence of stochastic disturbances, thereby maintaining performance and safety in a more uncertain environment.

Theorem 1. Consider the stochastic nonlinear system (6) satisfying Assumptions 1–4, along with the controller (7). Let us assume that the noise is also measured and collected as (13). Then, the largest ellipsoid $\mathcal{E}(P, 1)$ contained within the polyhedral set $\mathcal{S}(F, g)$ is λ -contractive for the closed-loop system if and only if there exist a decision variable G_2 and non-negative matrices P and Y such that

$$\min_{P, Y, G_2} -\log \det(P) \quad (21)$$

$$\text{s.t. } Z_0 Y = \begin{bmatrix} P \\ 0_{(n_z-n) \times n} \end{bmatrix}, \quad (22)$$

$$Z_0 G_2 = \begin{bmatrix} 0_{n \times (n_z-n)} \\ I_{n_z-n} \end{bmatrix}, \quad (23)$$

$$\begin{bmatrix} P & ((X_1 - W_0)Y)^T \\ (*) & \lambda P \end{bmatrix} \succeq 0, \quad (24)$$

$$\begin{bmatrix} P & P F_j^T \\ (*) & g_j^2 \end{bmatrix} \succeq 0, \quad \forall j = 1, \dots, q, \quad (25)$$

$$P \succeq 0, \quad (26)$$

$$(X_1 - W_0)G_2 = 0. \quad (27)$$

Then, Problem 1 is solved, and the safe control gain is obtained by

$$K = U_0 [Y P^{-1} \ G_2]. \quad (28)$$

Proof. First, we start with the contractivity condition. According to Definition 5 and condition (27) (which is considered to cancel the effect of system nonlinearities), the expected value of the next state is $\mathbb{E}[x_{t+1}] = \bar{x}_{t+1} = X_1 G_1 x_t$. Hence, the set $\mathcal{E}(P, 1)$ contained in the polyhedral admissible set $\mathcal{S}$ is λ -contractive in expectation, if the following holds

$$\bar{x}_{t+1}^T P^{-1} \bar{x}_{t+1} \leq \lambda \bar{x}_t^T P^{-1} \bar{x}_t \quad (29)$$

or

$$(X_1 G_1)^T P^{-1} X_1 G_1 \leq \lambda P^{-1} \quad (30)$$

Define $Y = G_1 P$. Applying the Schur complement to (30) yields (24). Conditions (22) and (23) are also directly obtained in terms of (14).

We now establish criteria for ensuring the contractive set is encompassed by the admissible set. The inclusion of the ellipsoid $\mathcal{E}(P, 1)$ within the polytope $\mathcal{S}$ occurs if and only if (Boyd et al., 1994)

$$\max\{F_j x | x \in \mathcal{E}(P, 1)\} \leq g_j \quad \forall j = 1, \dots, q, \quad (31)$$

which corresponds to (25). This stems from the requirement that the ellipsoidal set remains inside the polytope, indicated by

$$F_j x \leq g_j \quad \forall j = 1, \dots, q. \quad (32)$$

When (32) is multiplied by its transpose, the resulting equation is

$$F_j x x^T F_j^T \leq g_j^2 \quad \forall j = 1, \dots, q. \quad (33)$$

Considering the definition of ellipsoidal sets, where $x x^T \preceq P$, the inequality (33) transforms into

$$F_j P F_j^T \leq g_j^2 \quad \forall j = 1, \dots, q. \quad (34)$$

Applying the Schur complement to (34) leads to the formulation of the constraint (25). This concludes the proof. $\square$

Building on the initial framework where noise measurements were assumed available, the discussion now advances by considering a more practical scenario where such measurements are not accessible. This leads to Theorem 2, where the focus shifts to designing a safe nonlinear controller under the constraint of having no direct information about the noise affecting the system. To compensate for this lack of data and to ensure the continued safety of the system, the design strategy transitions towards an optimization objective that seeks to minimize the variance of the closed-loop system. This approach is primarily aimed at reducing the impact of noise, thereby decreasing the risk of safety violations and enhancing the system's resilience to disturbances in uncertain conditions. Additionally, unlike Theorem 1, which considers the effects of nonlinearities as needing to be completely canceled, Theorem 2 aims to minimize their residual impact. This relaxes the restrictive assumption of complete cancellation and allows for a more feasible implementation of the control strategy under practical conditions.

Theorem 2. Consider the stochastic nonlinear system (6), which satisfies Assumptions 1–4, and is controlled by (7). Assume further that noise measurements are unavailable. In this scenario, the largest ellipsoid $\mathcal{E}(P, 1)$, which fits within the polyhedral set $\mathcal{S}(F, g)$, is λ -contractive for the closed-loop system if and only if there exist a decision variable G_2 , non-negative matrices P , Y , and M , and a positive scalar μ such that

$$\min_{P, Y, G_2, M, \mu} -\beta_1 \log \det(P) + \beta_2 \mu + \beta_3 V \quad (35)$$

$$\text{s.t. } Z_0 Y = \begin{bmatrix} P \\ 0_{(n_z-n) \times n} \end{bmatrix}, \quad (36)$$

$$Z_0 G_2 = \begin{bmatrix} 0_{n \times (n_z-n)} \\ I_{n_z-n} \end{bmatrix}, \quad (37)$$

$$\begin{bmatrix} P & (X_1 Y)^T \\ (*) & \lambda P \end{bmatrix} \succeq 0, \quad (38)$$

$$\begin{bmatrix} P & P F_j^T \\ (*) & g_j^2 \end{bmatrix} \succeq 0, \quad \forall j = 1, \dots, q, \quad (39)$$

$$\begin{bmatrix} M & Y \\ (*) & P \end{bmatrix} \succeq 0, \quad (40)$$

$$P \succeq 0, \quad (41)$$

$$\|X_1 G_2\| \leq \mu, \quad (42)$$

$$\mu > 0. \quad (43)$$

where $V = \text{Tr}(M) + \eta_1 \|G_1\|_2^2 + \eta_2 \|G_2\|_2^2$ and β_i for $i = 1, 2, 3$ are some user-defined weighting constants for the objective function. Then, Problem 1 is solved, and the safe control gain is derived by

$$K = U_0 [Y P^{-1} \ G_2]. \quad (44)$$

Proof. The nominal value of the next state is computed as

$$\bar{x}_{t+1} = X_1 G_1 x_t + X_1 G_2 S(x_t). \quad (45)$$

For the random variable $\tilde{x}_{t+1} = x_{t+1} - \bar{x}_{t+1} = -W_0 G_1 x_t - W_0 G_2 S(x_t) + w(t)$, its covariance satisfies

$$\begin{aligned}
\mathbb{E}[\tilde{x}_{t+1}\tilde{x}_{t+1}^\top] &= \mathbb{E}[W_0G_1x_t x_t^\top G_1^\top W_0^\top] \\
&+ 2\mathbb{E}[W_0G_1x_t S(x_t)^\top G_2^\top W_0^\top] \\
&+ \mathbb{E}[W_0G_2S(x_t)S(x_t)^\top G_2^\top W_0^\top] + \Sigma \\
&\preceq \mathbb{E}[W_0G_1PG_1^\top W_0^\top] + 2\mathbb{E}[W_0G_1x_t S(x_t)^\top G_2^\top W_0^\top] \\
&+ \mathbb{E}[W_0G_2S(x_t)S(x_t)^\top G_2^\top W_0^\top] + \Sigma \quad (46)
\end{aligned}$$

where the inequality is derived from the observation that $x_t x_t^\top \preceq P$, given that $x_t^\top P^{-1} x_t \leq 1$.

Now, according to Lemma 1 and leveraging the Gaussian characteristics of W_0 , the expectation of the product involving Gaussian noise matrices simplifies due to independence as follows

$$E[W_0H_1W_0^\top] = \text{Tr}(H_1)\Sigma \quad (47)$$

where $H_1 = G_1xS(x)^\top G_2^\top$ is the deterministic part of the matrix product. Therefore, utilizing the properties of the trace function for matrices, it follows that

$$\text{Tr}(G_1xS(x)^\top G_2^\top) \leq \|G_1\|_2 \|xS(x)^\top\|_F \|G_2^\top\|_2 \quad (48)$$

To find the most adverse value of $\|xS(x)^\top\|_F$ over the polyhedral set, one has the following optimization problem

$$\max \|xS(x)^\top\|_F \quad (49)$$

$$\text{s.t. } F_j x \leq g_j \quad (50)$$

Using the approach detailed in (Ahmadi, 2016), this optimization problem can be equivalently formulated as

$$\max \eta_1 \quad (51)$$

$$\text{s.t. } -\|xS(x)^\top\|_F - \eta_1 \geq 0, \quad (52)$$

$$F_j x \leq g_j, \quad \forall j = 1, \dots, q. \quad (53)$$

The objective is to find the largest η_1 such that $-\|xS(x)^\top\|_F - \eta_1$ is nonnegative for all x satisfying the constraints $F_j x \leq g_j \forall j = 1, \dots, q$, which can be formulated as the following sum of squares (SOS) problem

$$\max_{\tau_j, \eta_1} \eta_1 \quad (54)$$

$$\text{s.t. } s_o(x) - \left((\|xS(x)^\top\|_F + \eta_1) + \sum_j \tau_j (g_j - F_j x) \right) \text{ is SOS}, \quad (55)$$

$$\tau_j \geq 0 \quad \forall j = 1, \dots, q. \quad (56)$$

In this formulation, $s_o(x)$ is an SOS polynomial, and $\tau_j \forall j = 1, \dots, q$ are decision variables.

Given the bound on the Frobenius norm of $xS(x)^\top$, one gets

$$\text{Tr}(G_1xS(x)^\top G_2^\top) \leq \|G_1\|_2 \eta_1 \|G_2\|_2 \quad (57)$$

According to this result, and owing to the fact that the objective is to make the linear part of the system dominant, i.e., $\|G_2\| < \|G_1\|$, we derive the upper bound for the expected matrix

$$E[W_0G_1xS(x)^\top G_2^\top W_0^\top] \preceq \eta_1 \|G_1\|_2^2 \Sigma \quad (58)$$

In a similar manner, we have

$$E[W_0G_2S(x)S(x)^\top G_2^\top W_0^\top] \preceq \eta_2 \|G_2\|_2^2 \Sigma \quad (59)$$

where η_2 is the maximum value of the term $\|S(x)S(x)^\top\|_F$.

Hence, the inequality (46) becomes

$$\mathbb{E}[\tilde{x}_{t+1}\tilde{x}_{t+1}^\top] \preceq (\text{Tr}(G_1PG_1^\top) + \eta_1 \|G_1\|_2^2 + \eta_2 \|G_2\|_2^2) \Sigma + \Sigma \quad (60)$$

Now, consider a matrix M such that

$$G_1PG_1^\top = YP^{-1}Y^\top \preceq M. \quad (61)$$

which results in

$$\text{Tr}(YP^{-1}Y^\top) \leq \text{Tr}(M). \quad (62)$$

Applying the Schur complement on (61) yields (40). Also, (60) becomes

$$\mathbb{E}[\tilde{x}_{t+1}\tilde{x}_{t+1}^\top] \preceq (\text{Tr}(M) + \eta_1 \|G_1\|_2^2 + \eta_2 \|G_2\|_2^2) \Sigma + \Sigma \quad (63)$$

The impact of noise on the closed-loop system is reduced by minimizing the expression $V = \text{Tr}(M) + \eta_1 \|G_1\|_2^2 + \eta_2 \|G_2\|_2^2$. The remainder of the proof follows the same steps as in Theorem 1 and is therefore not repeated here. $\square$

4. SIMULATION

Simulations are carried out to demonstrate the efficiency of the proposed safe controller. The controller is applied to a discretized model of an inverted pendulum system, represented by the following equations

$$\begin{aligned}
x_{1,t+1} &= x_{1,t} + T_s x_{2,t} \\
x_{2,t+1} &= \frac{T_s g}{\ell} \sin(x_{1,t}) + \left(1 - \frac{T_s \mu}{m\ell^2}\right) x_{2,t} + \frac{T_s}{m\ell^2} u_t + w_t, \quad (64)
\end{aligned}$$

where $T_s = 0.1$ s denotes the sampling time, $m = 1$ kg represents the mass of the pendulum, $\ell = 1$ m is the length of the string, $g = 9.8$ m/s² stands for the gravitational acceleration, $\mu = 0.01$ is the coefficient of rotational friction, u is the control input, and w is the White Gaussian noise term introduced to simulate noises. The control parameter λ is selected as 0.98, and the dictionary of functions is $z(x_t) = [x_t^\top \sin(x_{1,t})]^\top$.

The states of the system and the corresponding phase plane trajectories are recorded. Figure 2 depicts the system states over time, showcasing the controller's capacity to stabilize the pendulum in the presence of noise. Figure 3 presents the phase plane of the system for different realizations of the Gaussian noise, which confirms the safety and stability of the closed-loop system under the designed nonlinear state-feedback control law. Each trajectory in the phase plane is plotted to illustrate the system's response in the presence of noise, showcasing the effectiveness of the safe controller in preserving the desired system behavior within predefined safety margins.

5. CONCLUSION

The paper presented a data-based control design for stochastic nonlinear systems through a double-part data-based closed-loop control representation. This included a linear component for λ -contractivity of the safe set to ensure its invariance, and a nonlinear component to mitigate the effects of nonlinear dynamics. The control gains were expressed through a closed-form data-based representation and optimized via a semidefinite programming problem (SDP) to enhance robustness against unmeasured noise and residual nonlinearities. The efficacy of this methodology was confirmed through simulations, demonstrating its ability to significantly enhance safety and performance in complex system environments. Future work will focus on generalizing the approach to non-input-affine systems, incorporating input constraints, and considering other noise distributions.

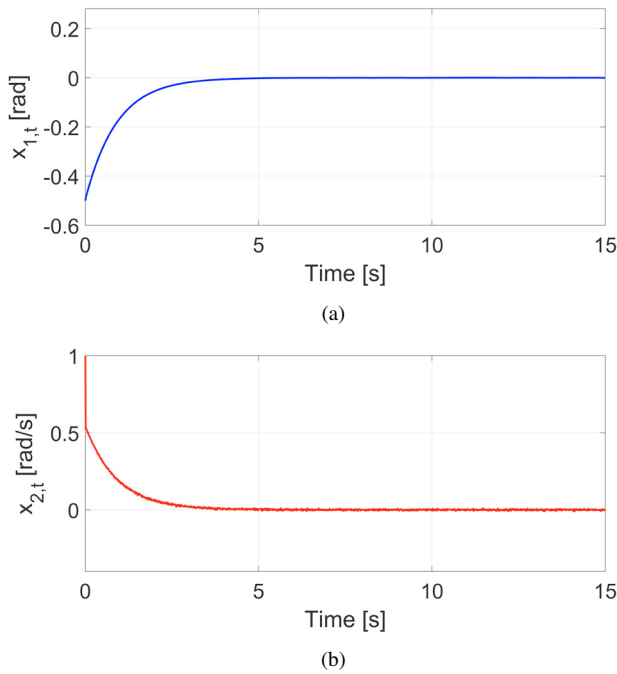


Fig. 2. Time history of the system states for one of the realizations of the Gaussian noise with $\Sigma = 0.00001I$ using the designed nonlinear safe controller.

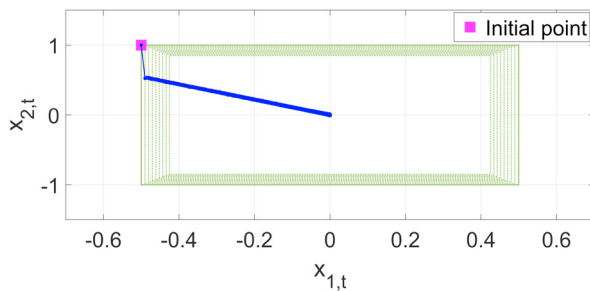


Fig. 3. Time history of the system trajectory for 100 different realizations of the Gaussian noise with $\Sigma = 0.00001I$ using the designed nonlinear safe controller.

REFERENCES

- Ahmadi, A.A. (2016). Sum of squares (SOS) techniques : An introduction. URL <https://api.semanticscholar.org/CorpusID:33703226>.
- Ahmadi, M., Israel, A., and Topcu, U. (2020). Safe controller synthesis for data-driven differential inclusions. *IEEE Transactions on Automatic Control*, 65(11), 4934–4940.
- Alshiekh, M., Bloem, R., Ehlers, R., Könighofer, B., Niekum, S., and Topcu, U. (2018). Safe reinforcement learning via shielding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2669–2678.
- Bisoffi, A., De Persis, C., and Tesi, P. (2020). Data-based guarantees of set invariance properties. *IFAC-PapersOnLine*, 53(2), 3953–3958.
- Blanchini, F. and Miani, S. (2008). *Set-theoretic methods in control*, volume 78. Springer.
- Boyd, S., El Ghaoui, L., Feron, E., and Balakrishnan, V. (1994). *Linear matrix inequalities in system and control theory*. SIAM.
- Brunke, L., Greeff, M., Hall, A.W., Yuan, Z., Zhou, S., Panerati, J., and Schoellig, A.P. (2022). Safe learning in robotics:

- From learning-based control to safe reinforcement learning. *Annual Review of Control, Robotics, and Autonomous Systems*, 5, 411–444.
- Coppens, P., Schuurmans, M., and Patrinos, P. (2020). Data-driven distributionally robust LQR with multiplicative noise.
- De Persis, C., Rotulo, M., and Tesi, P. (2023). Learning controllers from data via approximate nonlinearity cancellation. *IEEE Transactions on Automatic Control*.
- Esmaeili, B., Salim, M., and Baradarannia, M. (2022). Pre-defined performance-based model-free adaptive fractional-order fast terminal sliding-mode control of MIMO nonlinear systems. *ISA Transactions*, 131, 108–123.
- Grandia, R., Taylor, A.J., Ames, A.D., and Hutter, M. (2021). Multi-layered safety for legged robots via control barrier functions and model predictive control. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 8352–8358. IEEE.
- Guo, M., De Persis, C., and Tesi, P. (2023). Learning control of second-order systems via nonlinearity cancellation. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, 3055–3060. IEEE.
- Hou, Z.S. and Wang, Z. (2013). From model-based control to data-driven control: Survey, classification and perspective. *Information Sciences*, 235, 3–35.
- Lan, Y. and Mezić, I. (2013). Linearization in the large of nonlinear systems and koopman operator spectrum. *Physica D: Nonlinear Phenomena*, 242(1), 42–53.
- Li, S. and Bastani, O. (2020). Robust model predictive shielding for safe reinforcement learning with stochastic dynamics. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 7166–7172. IEEE.
- Mazouchi, M., Nagesh Rao, S., and Modares, H. (2021). Conflict-aware safe reinforcement learning: A meta-cognitive learning framework. *IEEE/CAA Journal of Automatica Sinica*, 9(3), 466–481.
- Modares, A., Sadati, N., Esmaeili, B., Yaghmaie, F.A., and Modares, H. (2023). Safe reinforcement learning via a model-free safety certifier. *IEEE Transactions on Neural Networks and Learning Systems*.
- Modares, H. (2023). Data-driven safe control of uncertain linear systems under aleatory uncertainty. *IEEE Transactions on Automatic Control*, 1–8. doi:10.1109/TAC.2023.3267019.
- Reis, M.F., Aguiar, A.P., and Tabuada, P. (2020). Control barrier function-based quadratic programs introduce undesirable asymptotically stable equilibria. *IEEE Control Systems Letters*, 5(2), 731–736.
- Zanon, M. and Gros, S. (2020). Safe reinforcement learning using robust MPC. *IEEE Transactions on Automatic Control*, 66(8), 3638–3652.
- Zhang, L., Zhang, R., Wu, T., Weng, R., Han, M., and Zhao, Y. (2021). Safe reinforcement learning with stability guarantee for motion planning of autonomous vehicles. *IEEE Transactions on Neural Networks and Learning Systems*, 32(12), 5435–5444.