IEEE Control Systems Letters paper presented at
2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Data-Driven Synthesis of Safety Controllers via Multiple Control Barrier Certificates

Ameneh Nejati, *Student Member, IEEE*, and Majid Zamani, *Senior Member, IEEE*

Abstract— **This work proposes a data-driven framework to synthesize safety controllers for nonlinear systems with finite input sets and unknown mathematical models. The proposed scheme leverages new notions of *multiple control barrier certificates (M-CBC)* and provides controllers ensuring the safety of systems with confidence 1. While there may not exist a *common* control barrier certificate with a fixed template, our proposed technique adaptively partitions the state set to potentially find M-CBC of the same template for different regions. In the proposed data-driven framework, we first cast our proposed conditions of M-CBC as a robust optimization program (ROP). Given that the unknown model appears in some of the constraints of the ROP, we propose a sampling approach for collecting data and provide a scenario optimization program (SOP) associated with the proposed ROP. We solve the resulted SOP and construct M-CBC together with safety controllers for the unknown system with 100% correctness guarantee. We apply our results to a nonlinear jet engine compressor with unknown dynamics to illustrate the efficacy of our data-driven approach. In the case study, we show that while there exists no *common* polynomial-type control barrier certificate of a given degree, there exist polynomial-type M-CBC of the same degree by partitioning the state set to different regions.**

I. Introduction

Formal methods have become popular, over the past two decades, for providing formal analyses over complex dynamical systems. In general, providing formal verification and controller synthesis frameworks for complex systems to enforce high-level logic properties, *e.g.,* those expressed as linear temporal logic (LTL) formulae [1], is very challenging. This is mainly due to (i) continuity of state sets, (ii) dealing with complex logic requirements, and (iii) lack of closed-form mathematical models in many real-world applications.

To alleviate the aforementioned difficulties, one promising approach, proposed initially in [2], [3], is to employ *barrier certificates* as a discretization-free approach for the formal verification and controller synthesis of dynamical systems. In particular, barrier certificates are some Lyapunov-like functions whose level sets separate an unsafe region from system's trajectories originating from a given set. As a result, the existence of such a function provides a (probabilistic) safety certificate for the system. Over the past decade, barrier certificates have been extensively utilized for formal verification and controller synthesis of non-stochastic [4], [5] and stochastic systems [6]–[8], to name a few.

Unfortunately, the above-mentioned results require knowing the precise models to provide corresponding analyses. Accordingly, one cannot leverage those techniques when the model of the system is unknown, which is the case in many real-life applications. To tackle this difficulty, there have been some *indirect* data-driven techniques based on systems identification to approximate underlying dynamics followed by model-based analysis approaches (see [9]–[11]). However, those techniques are mainly limited to linear or some particular classes of nonlinear systems and acquiring a precise model for complex systems is generally computationally expensive (see e.g., [12, and references herein]). Due to the underlying difficulty, the main goal of this work is to develop a *direct* data-driven technique to bypass the system identification phase and directly construct a barrier certificate by collecting data from trajectories of the unknown system.

The main contribution of this work is to propose a data-driven technique to synthesize safety controllers for nonlinear systems with finite input sets and unknown mathematical models. In general, there may not exist any *common* control barrier certificate of a fixed template together with its corresponding controller for the whole range of the state set to enforce the safety of the system. In this work, we propose a new technique to adaptively partition the state set and construct *multiple* control barrier certificates (M-CBC) for different regions. We first cast our new conditions of M-CBC as a robust optimization program (ROP). Given that the unknown model appears in some of the constraints of the ROP, we provide a scenario optimization program (SOP) associated with the ROP by proposing a sampling approach and collecting data from the system. By solving the acquired SOP, we construct M-CBC together with safety controllers for the unknown system with 100% correctness guarantee. We show the efficacy of our proposed results over a nonlinear jet engine compressor with unknown models.

Related Work. In the past few years, several studies have been performed on the formal analysis of unknown dynamical systems via *direct* data-driven approaches. Existing results include: data-driven learning of control laws ensuring stability of nonlinear polynomial-type models [13]; stability verification of unknown switched linear systems via data [14]; data-driven synthesis of state-feedback controllers to make a compact polyhedral set including the origin invariant [15], [16]; and data-driven approaches for the verification and controller synthesis of unknown dynamical systems via control barrier certificates [17]–[19], to name a few.

It is worth mentioning that the results in [15], [16] provide data-driven synthesis of state-feedback controllers to make

This work was supported in part by the NSF under grants CNS-2145184, CNS-1952223, and A22-0123-S003.

A. Nejati is with the Department of Computer Science, LMU Munich, Germany. Majid Zamani is with the Department of Computer Science, University of Colorado Boulder, USA. Email: amy.nejati@lmu.de, majid.zamani@colorado.edu.

the whole state set invariant. In comparison, we propose here a less conservative approach by adaptively partitioning the state set and constructing multiple CBC together with corresponding safety controllers for different regions. The data-driven results in [17] are only tailored to nonlinear *polynomial-type* systems, whereas our data-driven controller synthesis framework here is applicable to *any class of nonlinear systems* which are locally Lipschitz continuous. Note that the data-driven approaches in [18], [19] come with a probabilistic confidence level, whereas we propose here a deterministic sampling technique to construct M-CBC from data together with a safety controller with 100% correctness guarantee. As another pivotal difference, the proposed results in [17]- [19] provide a *common* control barrier certificate for unknown systems, whereas we develop here a new notion of *multiple control barrier certificates* which is more general.

II. Discrete-Time Nonlinear Control Systems

A. Notation

We denote sets of nonnegative and positive integers by $\mathbb{N} := \{0, 1, 2, \ldots\}$ and $\mathbb{N}^+ := \{1, 2, 3, \ldots\}$, respectively. Moreover, symbols $\mathbb{R}$, $\mathbb{R}^+$, and $\mathbb{R}_0^+$ denote, respectively, sets of real, positive, and nonnegative real numbers. Symbols $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ are used to denote, respectively, an n-dimensional Euclidean space and the space of real matrices with n rows and m columns. The Euclidean norm of $x \in \mathbb{R}^n$ is denoted by $\|x\|$. For any symmetric matrix $P \in \mathbb{R}^{m\times n}$, we have $\|P\| := \sqrt{\lambda_{\max}(P^\top P)}$, where $\lambda_{\max}(\cdot)$ is the maximum eigenvalue. Given N column vectors $x_i \in \mathbb{R}^{n_i}$, $n_i \in \mathbb{N}^+$, and $i \in \{1, \ldots, N\}$, $x = [x_1; \ldots; x_N]$ denotes a column vector of the dimension $\sum_i n_i$.

B. Discrete-Time Nonlinear Control Systems

Now, we define discrete-time nonlinear control systems (dt-NCS) as the underlying model in this work.

Definition 2.1: A discrete-time nonlinear control system (dt-NCS) is characterized by

$$\Sigma : x(k+1) = f(x(k), \nu(k)), \quad k \in \mathbb{N}, \tag{1}$$

where:

- $X \subseteq \mathbb{R}^n$ is the state set;
- $U = \{u_1, u_2, \ldots, u_m\}$ with $u_i \in \mathbb{R}^{\bar{m}}, i \in \{1, \ldots, m\}$, is the finite input set, and $\nu : \mathbb{N} \to U$ is the input sequence;
- $f : X \times U \to X$ is the transition map which is assumed to be *unknown*.

We denote the state trajectory of dt-NCS at time $k \in \mathbb{N}$, under the input trajectory $\nu(\cdot)$, and starting from x_0 by $x_{x_0,\nu}(k)$.

In the next section, we provide a new notion of multiple control barrier certificates for discrete-time nonlinear control systems in (1).

III. Multiple Control Barrier Certificates

In general, constructing common control barrier certificates for the whole range of the state set is conservative. To alleviate this conservatism, we adaptively partition the state set of dt-NCS in (1) and propose a new notion of multiple control barrier certificates (M-CBC) to be constructed in different regions, as formalized in the next definition.

Definition 3.1: Consider a dt-NCS Σ in (1) and let $X = \cup_{i=1}^N X_i$, with $X_i \subseteq X$ being a partition element where $X_i \cap X_j = \varnothing$ for any $i \neq j$. Consider $X_0, X_u \subseteq X$ as initial and unsafe sets of dt-NCS, respectively. A collection of functions $\mathcal{B}_i : X_i \to \mathbb{R}$ is called multiple control barrier certificates (M-CBC) for Σ with respect to an input set $U = \{u_1, u_2, \ldots, u_m\}$, if $\forall i \in \{1, \ldots, N\}$:

$$\forall x \in X_i \cap X_0, \qquad \mathcal{B}_i(x) \leq \gamma_i, \tag{2}$$

$$\forall x \in X_i \cap X_u, \qquad \mathcal{B}_i(x) \geq \lambda_i, \tag{3}$$

$$\forall j \in \{1, \ldots, N\}, \qquad \lambda_i > \gamma_j, \text{ with } \lambda_i, \gamma_j \in \mathbb{R}, \tag{4}$$

and $\forall x \in X_i, \exists u \in U, \exists j \in \{1, \ldots, N\}$, such that:

$$f(x, u) \in X_j \implies \mathcal{B}_j(f(x, u)) \leq \mathcal{B}_i(x). \tag{5}$$

It is worth emphasizing that condition (5) does not take into account all pairs (i, j) since the state trajectory of the system may not necessarily enter other cells in one-step transition. Note that one can construct a set-valued controller $\mathcal{C} : X \to 2^U$ based on $\mathcal{B}_i(x)$ as $\mathcal{C}(x) = \{u \in U \,|\, \mathcal{B}_j(f(x, u)) \leq \mathcal{B}_i(x), \text{for some } j \in \{1, \ldots, N\} \text{ where } f(x, u) \in X_j\}$ for any $x \in X_i$. Since f is unknown, we explain in Section IV how this controller is designed based on data collected from the system's trajectories.

In the next definition, we present the main safety problem for dt-NCS Σ.

Definition 3.2: Consider a dt-NCS Σ. Given a safety specification $\varphi = (X_0, X_u)$, where $X_0, X_u \subseteq X$ with $X_0 \cap X_u = \emptyset$, Σ is called safe, denoted by $\Sigma \models \varphi$, if all trajectories of Σ started from the initial set $X_0 \subseteq X$ never reach the unsafe set $X_u \subseteq X$.

The next theorem, inspired by [2, Theorem 3], shows the usefulness of M-CBC in Definition 3.1 for ensuring the safety of dt-NCS as in Definition 3.2.

Theorem 3.3: Consider a dt-NCS Σ as in (1) with a partition over its state set as $X = \cup_{i=1}^N X_i$. Suppose a collection of functions $\mathcal{B}_i : X_i \to \mathbb{R}$, $i \in \{1, \ldots, N\}$, is M-CBC for Σ as in Definition 3.1. Then one has $x_{x_0,\nu}(k) \notin X_u$ for any $x_0 \in X_0$ and any $k \in \mathbb{N}$ under the input trajectory $\nu(k) \in \mathcal{C}(x(k))$, $\forall k \in \mathbb{N}$.

Proof: We show the proof by contradiction. Assume there exists a collection of functions $\mathcal{B}_i$ satisfying conditions (2)-(5). Suppose $x_{x_0,\nu}$ starts at some $x_0 \in X_i \cap X_0$ for some $i \in \{1, \ldots, N\}$ and reaches $X_j \cap X_u$ for some $j \in \{1, \ldots, N\}$. According to (2)-(3), we have $\mathcal{B}_i(x(0)) \leq \gamma_i$ and $\mathcal{B}_j(x(k)) \geq \lambda_j$ for some $k \in \mathbb{N}$. Since $\mathcal{B}_i$ is M-CBC and using (5), one can recursively show $\lambda_j \leq \mathcal{B}_j(x(k)) \leq \mathcal{B}_i(x(0)) \leq \gamma_i$. This contradicts condition $\lambda_j > \gamma_i$ in (4), which completes the proof. ■

Remark 3.4: Note that in order to ensure the safety of the system via the results of Theorem 3.3, one can start partitioning the state set with an arbitrary N and then search for M-CBC satisfying conditions (2)-(5). If the required conditions for some cells are not satisfied, those particular cells can be partitioned more such that conditions (2)-(5) may eventually be fulfilled.

IV. Data-Driven Framework

In this section, we aim at constructing M-CBC using data collected from trajectories of systems. In our data-driven setting, we fix the structure of M-CBC as $\mathcal{B}_i(q_i, x) = \sum_{l=1}^{z_i} q_i^l p_i^l(x)$ with user-defined (potentially nonlinear) basis functions p_i^l and unknown coefficients $q_i = [q_i^1; \dots; q_i^{z_i}] \in \mathbb{R}^{z_i}$. It is worth mentioning that basis functions p_i^l can be considered to have any arbitrary form, *e.g.,* they can be monomials over x if one is interested in the polynomial-type M-CBC.

To fulfill conditions (2)-(5) in Definition 3.1, we recast our problem as the following robust optimization program (ROP):

$$\min_{[d_1;\dots;d_N;\eta]} \quad \eta,$$

$$\begin{aligned}
\text{s.t.}\ & \forall i \in \{1,\dots,N\}: \\
& \forall x \in X_i \cap X_0, \mathcal{B}_i(q_i, x) - \gamma_i \leq \eta_i, && \text{(6a)}\\
& \forall x \in X_i \cap X_u, -\mathcal{B}_i(q_i, x) + \lambda_i \leq \eta_i, && \text{(6b)}\\
& \forall j \in \{1,\dots,N\}, \gamma_j - \lambda_i \leq \eta_i && \text{(6c)}\\
& \forall x \in X_i, \exists u \in U, \exists j \in \{1,\dots,N\}: \\
& f(x,u) \in X_j \Longrightarrow \mathcal{B}_j(q_j, f(x,u)) - \mathcal{B}_i(q_i, x) \leq \eta_i, && \text{(6d)}\\
& \eta_i \leq \eta, && \text{(6e)}\\
& d_i = [\gamma_i; \lambda_i; q_i^1; \dots; q_i^{z_i}; \eta_i], \gamma_i, \lambda_i, q_i^l, \eta_i, \eta \in \mathbb{R}.
\end{aligned}$$

If $\eta \leq 0$, a solution to the ROP ensures conditions (2)-(5) in Definition 3.1 are satisfied.

To solve the proposed ROP in (6), one faces two major challenges. First, the ROP in (6) has infinitely many constraints since the state set of dt-NCS is continuous (*i.e.,* $x \in X_i$). In addition, the map f is required for solving the ROP, which is unknown in this work. To tackle those challenges, we aim at developing a data-driven scheme for the construction of M-CBC without solving the ROP in (6). To do so, we first collect M sampled data within X, denoted by

$$\big\{(\hat{x}_r, f(\hat{x}_r, u_s)) \,\big|\, r \in \{1,\dots,M\}, s \in \{1,\dots,m\}\big\}. \quad (7)$$

We then consider a ball X_r around each sample $\hat{x}_r$ with radius ε such that $X \subseteq \cup_{r=1}^{M} \mathsf{X}_r$ and

$$\|x - \hat{x}_r\| \leq \varepsilon, \qquad \forall x \in \mathsf{X}_r. \quad (8)$$

We now propose the following scenario optimization program (SOP), associated with the ROP in (6):

$$\min_{[d_1;\dots;d_N;\eta]} \quad \eta,$$

$$\begin{aligned}
\text{s.t.}\ & \forall i \in \{1,\dots,N\}, \forall r \in \{1,\dots,M\}: \\
& \forall \hat{x}_r \in X_i \cap X_0, \mathcal{B}_i(q_i, \hat{x}_r) - \gamma_i \leq \eta_i, && \text{(9a)}\\
& \forall \hat{x}_r \in X_i \cap X_u, -\mathcal{B}_i(q_i, \hat{x}_r) + \lambda_i \leq \eta_i, && \text{(9b)}\\
& \forall j \in \{1,\dots,N\}, \gamma_j - \lambda_i \leq \eta_i && \text{(9c)}\\
& \forall \hat{x}_r \in X_i, \exists u \in U, \exists j \in \{1,\dots,N\}: \\
& f(\hat{x}_r,u) \in X_j \Longrightarrow \mathcal{B}_j(q_j, f(\hat{x}_r,u)) - \mathcal{B}_i(q_i, \hat{x}_r) \leq \eta_i, && \text{(9d)}\\
& \eta_i \leq \eta, && \text{(9e)}\\
& d_i = [\gamma_i; \lambda_i; q_i^1; \dots; q_i^{z_i}; \eta_i], \gamma_i, \lambda_i, q_i^l, \eta_i, \eta \in \mathbb{R}.
\end{aligned}$$

One can readily see that $f(\hat{x}_r, u)$ in (9d) is the transition of the unknown dt-NCS after one-step starting from $\hat{x}_r$ under input u.

Remark 4.1: Note that condition (9d) can be rewritten as a max-min constraint:

$$\max_{\hat{x}_r \in X_i} \min_{u \in U} \big(\mathcal{B}_j(q_j, f(\hat{x}_r, u)) - \mathcal{B}_i(q_i, \hat{x}_r)\big) \leq \eta_i, \quad (10)$$

for those $j \in \{1,\dots,N\}$, where $f(\hat{x}_r, u) \in X_j$. In general, an optimization problem with max-min constraints is equal to a collection of optimization problems with inequality constraints. Solving such optimization problem could be potentially expensive due to having a large collection. Therefore, we employ the proposed approach in [20] and convert this condition into nonlinear programming in which the condition is a single inequality constraint as the following, $\forall i \in \{1,\dots,N\}, \forall r \in \{1,\dots,M\}$:

$$\sum_{s=1}^{m} \mu_s \big(\mathcal{B}_j(q_j, f(\hat{x}_r, u_s)) - \mathcal{B}_i(q_i, \hat{x}_r) - \eta_i\big) \leq 0, \quad (11)$$

for those $j \in \{1,\dots,N\}$, such that $f(\hat{x}_r, u_s) \in X_j$, and where $\sum_{s=1}^{m} \mu_s = 1$, $\mu_s \in \mathbb{R}_0^+$. The max-min constraint in (10) is satisfied if and only if the single inequality constraint in (11) is fulfilled [20, Proposition 2.1]. The resulting optimization program can then be solved using available software tools such as NPSOL [21].

V. Data-Driven Construction of M-CBC

In this section, we aim at solving the proposed SOP in (9) and constructing M-CBC for unknown dt-NCS with a guaranteed confidence of 1. To do so, we first raise the following assumption.

Assumption 1: Suppose $\mathcal{B}_i(q_i, x)$ is Lipschitz continuous with respect to x with a Lipschitz constant $\mathscr{L}_1$, for any $i \in \{1,\dots,N\}$, and $\mathcal{B}_j(q_j, f(x,u)) - \mathcal{B}_i(q_i, x)$ in (6d) is Lipschitz continuous with respect to x with a Lipschitz constant $\mathscr{L}_2$, for any $i, j \in \{1,\dots,N\}$, and any input $u \in U$.

Under Assumption 1, the next theorem provides a data-driven construction scheme for M-CBC over unknown dt-NCS with a certified confidence of 1.

Theorem 5.1: Given an unknown dt-NCS in (1), let Assumption 1 hold. Suppose the SOP in (9) is solved with $M \times m$ sampled data as in (7) with an optimal value η_M^* and solution $d_i^* = [\gamma_i^*; \lambda_i^*; q_i^{1*}; \dots; q_i^{z_i*}; \eta_i^*], \forall i \in \{1,\dots,N\}$. If

$$\eta_M^* + \mathscr{L}\varepsilon \leq 0, \quad (12)$$

with $\mathscr{L} = \max\{\mathscr{L}_1, \mathscr{L}_2\}$, then the constructed $\{\mathcal{B}_1, \dots, \mathcal{B}_N\}$ via solving SOP in (9) are M-CBC for unknown dt-NCS with a confidence of 1. Hence, there exists a set-valued controller $\mathcal{C}$ under which the unknown dt-NCS is safe in the sense of Theorem 3.3.

Proof: We first show that, under condition (12), the constructed $\mathcal{B}_i$ via solving SOP in (9) satisfy (5) for the whole range of X_i, *i.e.,* for any $x \in X_i$, there exists $u \in U$ such that:

$$\mathcal{B}_j(q_j, f(x,u)) - \mathcal{B}_i(q_i, x) \leq 0.$$

Note that according to (8), for any $x \in X_i$, there exists $\hat{x}_r \in \mathsf{X}_r$ such that x and $\hat{x}_r$ are ε-close, *i.e.,* $\|x - \hat{x}_r\| \leq \varepsilon$. One can readily observe from (9d) that for any $\hat{x}_r$, there exists a choice of $u \in U$, namely u^*, such that $\mathcal{B}_j(q_j, f(\hat{x}_r, u^*)) - \mathcal{B}_i(q_i, \hat{x}_r) \leq \eta_M^*$. Since $\mathcal{B}_j(q_j, f(x, u^*)) - \mathcal{B}_i(q_i, x)$ is Lipschitz continuous with respect to x with Lipschitz constant $\mathscr{L}_2$, we have, $\forall i \in \{1, \ldots, N\}, \forall r \in \{1, \ldots, M\}$:

$$\begin{aligned}&\mathcal{B}_j(q_j,f(x,u^*))-\mathcal{B}_i(q_i,x) = \mathcal{B}_j(q_j,f(x,u^*))-\mathcal{B}_i(q_i,x)\\ &-(\mathcal{B}_j(q_j,f(\hat{x}_r,u^*))-\mathcal{B}_i(q_i,\hat{x}_r))+(\mathcal{B}_j(q_j,f(\hat{x}_r,u^*))-\mathcal{B}_i(q_i,\hat{x}_r))\\ &\leq \mathscr{L}_2\|x-\hat{x}_r\| + \eta_M^* \leq \mathscr{L}\varepsilon + \eta_M^*.\end{aligned}$$

Since $\eta_M^* + \mathscr{L}\varepsilon \leq 0$, one can readily verify that for any $x \in X_i$, there exists $u \in U$ such that:

$$\mathcal{B}_j(q_j,f(x,u))-\mathcal{B}_i(q_i,x) \leq 0.$$

We now leverage a similar argument and show that, under condition (12), the constructed $\mathcal{B}_i$ via solving SOP in (9) satisfy (2) for any $x \in X_i \cap X_0$, as well. Since $\mathcal{B}_i(q_i, x)$ is Lipschitz continuous with Lipschitz constant $\mathscr{L}_1$ according to Assumption 1, and given that $\mathcal{B}_i(q_i, \hat{x}_r) - \gamma_i \leq \eta_M^*$ according to (9a), one has, $\forall i \in \{1, \ldots, N\}, \forall r \in \{1, \ldots, M\}$:

$$\begin{aligned}\mathcal{B}_i(q_i,x)-\gamma_i&=\mathcal{B}_i(q_i,x)-\gamma_i-(\mathcal{B}_i(q_i,\hat{x}_r)-\gamma_i)+(\mathcal{B}_i(q_i,\hat{x}_r)-\gamma_i)\\ &\leq \mathscr{L}_1\|x-\hat{x}_r\| + \eta_M^* \leq \mathscr{L}\varepsilon + \eta_M^*.\end{aligned}$$

Since $\eta_M^* + \mathscr{L}\varepsilon \leq 0$, one can readily verify that

$$\mathcal{B}_i(q_i,x)-\gamma_i \leq 0, \qquad \forall x \in X_i \cap X_0.$$

One can employ the same argument and show that the constructed $\mathcal{B}_i$ via solving SOP in (9) satisfy (3) for any $x \in X_i \cap X_u$, as well. Then the constructed $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$ via solving SOP in (9) are M-CBC for unknown dt-NCS in (1) with the confidence of 1, which concludes the proof. ■

Remark 5.2: In order to satisfy $X \subseteq \cup_{r=1}^{M} \mathsf{X}_r$ under condition (8), the number of samples M can be computed based on the parameter ε as $M = \frac{\text{Vol}(X)}{\varepsilon^n}$, where $\text{Vol}(\cdot)$ represents the volume of a set. To potentially reduce the number of samples required, one can begin by collecting samples using a larger value of ε to solve the SOP in (9). If condition (12) is not satisfied with the chosen (potentially large) ε, it is necessary to select a smaller ε and solve the SOP again. If the state set is manually gridded, condition (8) can be readily met by implementing a uniform gridding approach. When dealing with real data, one can consider a sufficiently large ε (worst-case scenario) that ensures the satisfaction of condition (8).

Remark 5.3: To the best of our knowledge, almost all data-driven approaches whose main goal is to certify with 100% correctness guarantee some properties over unknown systems via data suffer from the so-called, *sample complexity*: the number of data for providing final guarantees is exponential with respect to the dimension of the underlying system. This is the case also in our work and potential ways to mitigate this computational complexity are to employ either divide and conquer strategy (a.k.a. compositional techniques) or parallelization over SOP. We will defer these approaches for future work.

Remark 5.4: Note that reducing the value of ε does not necessarily guarantee the eventual satisfaction of $\eta_M^* + \mathscr{L}\varepsilon \leq 0$ and, hence, existence of M-CBC. This is primarily due to the fact that the existence of M-CBC in our work is only sufficient for the synthesis of a controller, but it is not a necessary condition.

The results of Theorem 5.1 ensure that there exists a set-valued controller $\mathcal{C}$ under which unknown dt-NCS is safe in the sense of Theorem 3.3. In particular, we construct the set-valued map $\mathcal{C}$ as follows, for any $x \in X_i$:

$$\begin{aligned}\mathcal{C}(x) :=\Big\{&u \in U \,\big|\, \mathcal{B}_j(q_j, f(\hat{x}_r, u)) - \mathcal{B}_i(q_i, \hat{x}_r) \leq \eta_i^*,\\ &\text{for some } j \in \{1, \ldots, N\} \text{ where } f(x,u) \in X_j\\ &\text{and } \exists r \in \{1, \ldots, M\} \text{ such that } \|x - \hat{x}_r\| \leq \varepsilon\Big\}.\end{aligned} \tag{13}$$

The set-valued map $\mathcal{C}$ for any $x \in X_i$ is not empty according to condition (8) and Remark 4.1.

Remark 5.5: The set-valued map $\mathcal{C}$ in (13) intuitively implies that after solving the SOP in (9) and acquiring the M-CBC, one can a-posteriori check condition (9d) for all sampled data using the obtained M-CBC, and construct and store safety controllers in the form of a *lookup table*. The constructed lookup table can then be used in runtime as follows: for any measurement of the system $x \in X$, one can find the nearest data point $\hat{x}_r$ such that $\|x - \hat{x}_r\| \leq \varepsilon$. Then corresponding control inputs valid for $\hat{x}_r$ are also valid inputs for x.

Remark 5.6: Note that our results offer a significant advantage over indirect approaches such as system identification. In particular, the direct data-driven technique we propose has the capability to provide safety guarantees for a broad range of nonlinear systems that exhibit Lipschitz continuity. In contrast, system identification approaches are primarily designed for linear systems or specific classes of nonlinear systems. Moreover, even if the underlying dynamics can be learned through identification techniques, it remains necessary to construct a barrier certificate for the acquired model. Consequently, the computational complexity arises at two levels: model identification and barrier certificate construction. Furthermore, it is important to note that many identification techniques learn an approximate model with a certain level of probabilistic confidence. In contrast, our data-driven results offer a 100% correctness guarantee, ensuring a confidence level of 1. For more detailed information regarding the distinctions between direct and indirect data-driven techniques, we refer the interested readers to [22].

In order to check condition (12) in Theorem 5.1, one needs to first compute $\mathscr{L}$. In the following, we employ the proposed results in [23] and provide the following algorithm to estimate $\mathscr{L}_1, \mathscr{L}_2$ using a finite number of data.

Note that one can employ the procedure of Algorithm 1 and similarly estimate $\mathscr{L}_1$ using a finite number of data by considering $g(\hat{x}_r) = \mathcal{B}_i(q_i, \hat{x}_r)$ in Step 5. Under Algorithm 1, the following lemma, borrowed from [23], ensures the convergence of the estimated $\mathscr{L}_1, \mathscr{L}_2$ to their actual

Algorithm 1 Estimation of $\mathscr{L}_2$ via data

```
Require: $N, m, \mathcal{B}_i, \mathcal{B}_j$
 1: Choose $\hat{M}, Z \in \mathbb{N}^+$ and $\alpha \in \mathbb{R}^+$
 2: for $i = 1 : N$
 3:     for $s = 1 : m$
 4:         Select $\hat{M}$ sampled pairs $(\hat{x}_r, \hat{x}'_r)$ from $X_i$ such
            that $\|\hat{x}_r - \hat{x}'_r\| \leq \alpha$ for any $r \in \{1, \ldots, \hat{M}\}$
 5:         Compute the slope $S_r$ as
                $S_r = \frac{\|g(\hat{x}_r) - g(\hat{x}'_r)\|}{\|\hat{x}_r - \hat{x}'_r\|}, \quad \forall r \in \{1, \ldots, \hat{M}\},$
            with $g(\hat{x}_r) = \mathcal{B}_j(q_j, f(\hat{x}_r, u_s)) - \mathcal{B}_i(q_i, \hat{x}_r)$, for
            some $j \in \{1, \ldots, N\}$ where $f(\hat{x}_r, u_s) \in X_j$,
            ($g(\hat{x}'_r)$ is computed similarly)
 6:         Compute the maximum slope as
                $\psi = \max\{S_1, \ldots, S_{\hat{M}}\}$
 7:         Repeat Steps 4-6 $Z$ times and acquire $\psi_1, \ldots, \psi_Z$
 8:         Apply Reverse Weibull distribution [23] to $\psi_1, \ldots,$
            $\psi_Z$, which gives us so-called location, scale, and
            shape parameters
 9:         The obtained *location parameter* is the estimated
            $\mathscr{L}_{ij}^s$
10:     end
11: end
Ensure: $\mathscr{L}_2 = \max_{ijs} \mathscr{L}_{ij}^s$
```

values in the limit.

Lemma 5.7: Under Algorithm 1, the estimated $\mathscr{L}_1, \mathscr{L}_2$ converge to their actual values if and only if α goes to zero and $\hat{M}, Z$ go to infinity.

Note that one can pick α very small and $\hat{M}, Z$ very big to get a precise approximation for $\mathscr{L}_1, \mathscr{L}_2$.

VI. Case Study: Jet Engine Compressor

To show the efficacy of our data-driven results, we apply our approach to the following discrete-time *nonlinear* jet engine compressor [24]:

$$\Sigma\colon \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} x_1(k) + (-x_2(k) - \frac{3}{2}x_1^2(k) - \frac{1}{2}x_1^3(k))\tau \\ x_2(k) + (x_1(k) - u(k))\tau \end{bmatrix},$$

where $x_1 = \Phi - 1, x_2 = \Psi - \Lambda - 2$, with Φ, Ψ, Λ being the mass flow, the pressure rise, and a constant, respectively. Moreover, $u \in U = \{-1, -0.9, -0.8, \ldots, 0.8, 0.9, 1\}$, and $\tau = 0.1$ as the sampling time. The regions of interest are $X = [-1, 1]^2, X_0 = [-0.6, 0.6] \times [-0.7, 0.7]$, and $X_u = [-0.9, 0.9] \times [-1, -0.8] \cup [-0.9, 0.9] \times [0.8, 1]$. The model is assumed to be unknown. We aim at constructing M-CBC via solving SOP in (9) with a confidence of 1 while synthesizing controllers $\mathcal{C}$ under which unknown dt-NCS remains in the safe set $X \backslash X_u$ according to Theorem 3.3.

We first fix a *common* CBC in the form of $\mathcal{B}(q, x) = q^1x_1^4 + q^2x_1^2 + q^3x_1^2x_2^2 + q^4x_2^2 + q^5x_2^4 + q^6$. We also fix $\varepsilon = 0.01$ and acquire $M = 40000$. We solve the SOP in (9) (for the common CBC) with $M = 40000$ and compute coefficients of a *common* CBC for the whole range of the state set, *i.e.,* $N = 1$, together with other decision variables as

$$\begin{aligned} \mathcal{B}(q, x) = {} & 0.002x_1^4 - 0.0014x_1^2 - 0.0023x_1^2x_2^2 + 0.0084x_2^2 \\ & - 0.0067x_2^4 + 0.4, \ \eta_M^* = 0.0008. \end{aligned}$$

The resulting common CBC cannot ensure the safety of the unknown jet engine compressor given that the optimal value of SOP is positive.

We now apply our proposed results by partitioning our regions of interest into two different regions, *i.e.,* $N = 2$, as: $X_1 = [-1, 0] \times [-1, 1], X_{0_1} = [-0.6, 0] \times [-0.7, 0.7]$, $X_{u_1} = [-0.9, 0] \times [-1, -0.8] \cup [-0.9, 0] \times [0.8, 1]$; $X_2 = [0, 1] \times [-1, 1], X_{0_2} = [0, 0.6] \times [-0.7, 0.7]$, $X_{u_2} = [0, 0.9] \times [-1, -0.8] \cup [0, 0.9] \times [0.8, 1]$. We now consider the structure of our multiple CBC as the common one: $\mathcal{B}_i(q_i, x) = q_i^1x_1^4 + q_i^2x_1^2 + q_i^3x_1^2x_2^2 + q_i^4x_2^2 + q_i^5x_2^4 + q_i^6, i \in \{1, 2\}$. We solve the SOP in (9) with $M = 40000$ and compute coefficients of the multiple CBC together with other decision variables in SOP for two partitions:

$$\begin{aligned} \text{Region 1: } \mathcal{B}_1(q_1, x) &= 0.002x_1^4 - 0.0025x_1^2 + 0.0037x_1^2x_2^2 \\ & \quad + 0.4x_2^2 - 0.1515x_2^4 + 0.4, \\ \gamma_1^* &= 0.5704, \lambda_1^* = 0.5830, \eta_1^* = -0.0127, \\ \text{Region 2: } \mathcal{B}_2(q_2, x) &= 0.002x_1^4 - 0.0318x_1^2 + 0.0507x_1^2x_2^2 \\ & \quad + 0.4x_2^2 - 0.1356x_2^4 + 0.3935, \\ \gamma_2^* &= 0.5708, \lambda_2^* = 0.5812, \eta_2^* = -0.0126, \end{aligned}$$

with $\eta_M^* = -0.0126$. We now employ Algorithm 1 and compute $\mathscr{L}_1 = 0.2623, \mathscr{L}_2 = 0.9314$. Since $\eta_M^* + \mathscr{L}\varepsilon = -33 \times 10^{-4} \leq 0$, according to Theorem 5.1, one can guarantee that there exists a controller $\mathcal{C}$ under which the system is safe.

Satisfaction of conditions (2) and (3) via constructed M-CBC from data is illustrated in Fig. 1. As can be observed, initial sets X_{0_1}, X_{0_2} are inside their corresponding level sets (*i.e.,* $\mathcal{B}_1(q_1, x) = \gamma_1, \mathcal{B}_2(q_2, x) = \gamma_2$) and unsafe sets X_{u_1}, X_{u_2} are outside their corresponding level sets (*i.e.,* $\mathcal{B}_1(q_1, x) = \lambda_1, \mathcal{B}_2(q_2, x) = \lambda_2$). We now construct the safety controller as a lookup table for all sampled data and apply it to the unknown jet engine system. The closed-loop state trajectory of the unknown jet engine under the synthesized controller is also depicted in Fig. 1. As can be observed, the trajectory of the unknown jet engine remains in the safe set under the synthesized controller, depicted in Fig. 2. It took 20 seconds for solving SOP in (9) with $M = 40000$ samples on a machine with Windows operating system (Intel i7-8665U CPU with 16 GB of RAM).

VII. Conclusion

The primary objective of this study was to develop a data-driven approach for constructing multiple control barrier certificates (M-CBC) from available data, aiming to ensure the safety of unknown discrete-time nonlinear systems with a certified confidence level of 1. In pursuit of this objective, we introduced a scenario optimization program (SOP) that effectively utilized data gathered from trajectories of unknown systems. By successfully solving the SOP, we

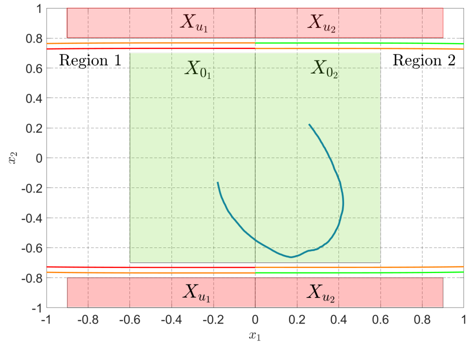

Fig. 1. Satisfaction of conditions (2)-(3). Green and pink boxes are initial and unsafe regions, respectively. Red and orange lines are initial and unsafe level sets of $\mathcal{B}_1$, respectively, for Region 1. Brown and green lines are initial and unsafe level sets of $\mathcal{B}_2$, respectively, for Region 2. Blue curve is closed-loop state trajectory of unknown jet engine starting from $x_0 = [-0.18; -0.16]$.

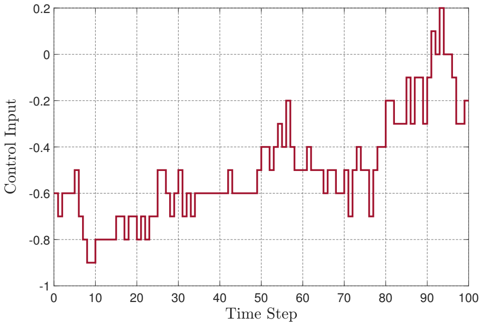

Fig. 2. A synthesized control input for the unknown jet engine within 100 time steps.

achieved the construction of the M-CBC, accompanied by its corresponding safety controller. Through evaluation on a nonlinear jet engine compressor with unknown models, we demonstrated the effectiveness and practicality of our data-driven approach. In fact, as the main contribution of our work, although a *common* control barrier certificate with a fixed template did not exist, our proposed technique adaptively partitioned the state set and found M-CBC of the same template for distinct regions. As a promising avenue for future research, we propose the extension of our data-driven technique to synthesize controllers capable of enforcing *more complex logic properties* for unknown nonlinear systems.

References

[1] A. Pnueli, “The temporal logic of programs,” in *Proceedings of the 18th Annual Symposium on Foundations of Computer Science*, 1977, pp. 46–57.

[2] S. Prajna and A. Jadbabaie, “Safety verification of hybrid systems using barrier certificates,” in *HSCC*, 2004, pp. 477–492.

[3] S. Prajna, A. Jadbabaie, and G. J. Pappas, “A framework for worst-case and stochastic safety verification using barrier certificates,” *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1415–1428, 2007.

[4] T. Wongpiromsarn, U. Topcu, and A. Lamperski, “Automata theory meets barrier certificates: Temporal logic verification of nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3344–3355, 2015.

[5] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, “Control barrier functions: Theory and applications,” in *Proceedings of the 18th European Control Conference (ECC)*, 2019, pp. 3420–3431.

[6] M. Ahmadi, B. Wu, H. Lin, and U. Topcu, “Privacy verification in POMDPs via barrier certificates,” in *Proceedings of the 57th IEEE Conference on Decision and Control (CDC)*, 2018, pp. 5610–5615.

[7] A. Nejati, S. Soudjani, and M. Zamani, “Compositional construction of control barrier functions for continuous-time stochastic hybrid systems,” *Automatica*, vol. 145, 2022.

[8] M. Anand, A. Lavaei, and M. Zamani, “Compositional synthesis of control barrier certificates for networks of stochastic systems against ω-regular specifications,” *Nonlinear Analysis: Hybrid Systems*, 2023.

[9] L. Wang, E. A. Theodorou, and M. Egerstedt, “Safe learning of quadrotor dynamics using barrier certificates,” in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2018, pp. 2460–2465.

[10] S. Sadraddini and C. Belta, “Formal guarantees in data-driven model identification and control synthesis,” in *Proceedings of the 21st International Conference on Hybrid Systems: Computation and Control*, 2018, pp. 147–156.

[11] L. Lindemann, H. Hu, A. Robey, H. Zhang, D. V. Dimarogonas, S. Tu, and N. Matni, “Learning hybrid control barrier functions from data,” in *Proceedings of the Conference on Robot Learning*, 2021, pp. 1351–1370.

[12] Z. Hou and Z. Wang, “From model-based control to data-driven control: Survey, classification and perspective,” *Information Sciences*, vol. 235, pp. 3–35, 2013.

[13] M. Guo, C. De Persis, and P. Tesi, “Learning control for polynomial systems using sum of squares relaxations,” in *59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 2436–2441.

[14] J. Kenanian, A. Balkan, R. M. Jungers, and P. Tabuada, “Data driven stability analysis of black-box switched linear systems,” *Automatica*, vol. 109, 2019.

[15] 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.

[16] ——, “Controller design for robust invariance from noisy data,” *IEEE Transactions on Automatic Control*, 2022.

[17] A. Nejati, B. Zhong, M. Caccamo, and M. Zamani, “Data-driven controller synthesis of unknown nonlinear polynomial systems via control barrier certificates,” in *Learning for Dynamics and Control Conference*, 2022, pp. 763–776.

[18] A. Nejati, A. Lavaei, P. Jagtap, S. Soudjani, and M. Zamani, “Formal verification of unknown discrete-and continuous-time systems: A data-driven approach,” *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 3011–3024, 2023.

[19] A. Salamati, A. Lavaei, S. Soudjani, and M. Zamani, “Data-driven verification and synthesis of stochastic systems through barrier certificates,” *arXiv: 2111.10330*, 2021.

[20] C. Kirjner-Neto and E. Polak, “On the conversion of optimization problems with max-min constraints to standard optimization problems,” *SIAM Journal on Optimization*, vol. 8, no. 4, pp. 887–915, 1998.

[21] P. Gill, W. Murray, M. Saunders, and M. Wright, “User’s guide for npsol: A fortran package for nonlinear programming. dept. of operation research,” Dept. of Operations Research, Stanford Univ. Technical Report SOL 86-2, Tech. Rep., 1986.

[22] F. Dörfler and I. Coulson, J.and Markovsky, “Bridging direct and indirect data-driven control formulations via regularizations and relaxations,” *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 883–897, 2022.

[23] G. Wood and B. Zhang, “Estimation of the Lipschitz constant of a function,” *Journal of Global Optimization*, vol. 8, no. 1, pp. 91–103, 1996.

[24] M. Krstic and P. V. Kokotovic, “Lean backstepping design for a jet engine compressor model,” in *Proceedings of International Conference on Control Applications*, 1995, pp. 1047–1052.