Data-Driven Safety-Certified Predictive Control for Linear Systems

Marjan Khaledi, Pouria Tooranjipour, and Bahare Kiumarsi

***Abstract*—A fully data-driven safe predictive control framework is presented for linear time-invariant (LTI) control systems. While model predictive control (MPC) is widely recognized for its ability to handle operational constraints, ensuring safety through maintaining the system within an invariant set is still an open challenge. In this letter, safety assurance is achieved through the integration of control barrier certificates (CBCs) in MPC. The behavioral systems theory is applied to obviate the system dynamics and consequently represent the MPC-CBC optimization using only input-state measurements. Furthermore, a data-driven maximal safe terminal set is constructed using the sum of squares (SOS) programming, surpassing the conventional sublevel sets of Lyapunov functions. This expansion of the terminal set leads to a significantly enlarged domain of attraction (DoA) for the MPC. The exponential stability and recursive feasibility of the proposed approach are rigorously proved by properly designing the terminal cost and the terminal set constraint. Finally, a numerical example is provided to illustrate the efficacy of the proposed MPC approach.**

***Index Terms*—Safety, MPC, data-driven control, control barrier certificates.**

I. Introduction

Ensuring the safety and performance of autonomous systems is key to their successful deployment. Model predictive control (MPC), as an optimal control technique, excels in performance and constraint handling [1] but faces two main limitations: ensuring safety within an invariant set [2] and reliance on precise system modeling [3]. This letter focuses on addressing these limitations.

To verify the safety of control systems, control barrier certificates (CBCs) have been widely developed as promising and effective tools [4]. CBCs certify safety by imposing set-invariant conditions under which the system's states remain in the safe set forever [5]. To simultaneously address safety and performance concerns, CBCs have been efficiently integrated into MPC [2], [6]. Notably, it has been demonstrated that the utilization of CBCs can reduce the prediction horizon in obstacle avoidance scenarios, leading to improvements in computational efficiency [2], albeit with a potential compromise on recursive feasibility. Although some efforts have been made to tackle this issue [6], they have primarily focused on achieving point-wise feasibility.

The feasible set of MPC is defined as the largest subset of state space for which there exists a control input satisfying all constraints. The invariance of this set is achieved by imposing an invariant terminal set. The size of the MPC feasible set is primarily influenced by two factors [7]: the size of the terminal region and the length of the prediction horizon. While increasing the prediction horizon enlarges the feasible set, it results in increased computational complexity, whereas obtaining a larger terminal set enlarges the feasible set without incurring an extra online cost. As a result, this letter proposes to enlarge the terminal set by utilizing the concept of CBCs [8]. This approach constructs a data-driven maximal safety-certified invariant set using offline data set through a sum-of-squares (SOS) program.

Statement of Contributions: The main contributions of this letter are threefold. First, by integrating the behavioral system theory [9] with the CBCs [4] within the MPC framework for the linear time-invariant (LTI) systems, we achieve two significant advancements: 1) relaxation of requiring the exact knowledge of the system dynamics compared to [2]. 2) a shortened prediction horizon, leading to enhanced computational efficiency. Second, through the application of CBCs, we construct a data-driven maximal safe terminal invariant set using SOS programming. According to [10], the CBC-based invariant set surpasses the size of conventional Lyapunov sublevel sets typically employed in standard MPC. Consequently, a more expansive terminal set corresponds to a broader DoA in the proposed MPC, leading to an extended feasible set and decreased online computational costs. Third, unlike [2], the terminal ingredients, i.e., the terminal cost and terminal set, are designed properly to ensure exponential stability and recursive feasibility, both of which are rigorously proven within this letter.

A. Notations

$I_{[a,b]}$ is the set of integers in the interval $[a, b]$. $x_{[0,N]}$ denotes the stacked vector $x_{[0,N]} = \left[x_0^T, \ldots, x_N^T\right]^T$. $\mathbb{R}_+$ denotes set of positive real numbers. $\mathbb{P}$ is a set of polynomials for $x \in \mathbb{R}^n$, and $\mathbb{P}^{SOS}$ is a set of SOS polynomials [11]. $p(x) \in \mathbb{P}^{SOS}$ if $p(x)$ can be rewritten as $p(x) = \sum_{i=1}^{i=m} p_i^2(x)$, where $p_i(x) \in \mathbb{P}$. $Vol(\mathcal{C})$ is

Manuscript received 14 September 2023; revised 16 November 2023; accepted 5 December 2023. Date of publication 11 December 2023; date of current version 28 December 2023. This work was supported in part by the Office of Naval Research under Award N00014-22-1-2159, and in part by the National Science Foundation under Award ECCS-2227311. Recommended by Senior Editor L. Zhang. *(Marjan Khaledi and Pouria Tooranjipour contributed equally to this work.) (Corresponding author: Bahare Kiumarsi.)*

The authors are with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824 USA (e-mail: khaledim@msu.edu; tooranji@msu.edu; kiumarsi@msu.edu).

Digital Object Identifier 10.1109/LCSYS.2023.3341346

the volume of $\mathcal{C}$. $\mathbb{N}_0$ denotes the set of whole numbers. The Kronecker product is denoted by $\otimes$, and vec(A) refers to the operation of stacking the columns of the matrix $A \in \mathbb{R}^{m\times n}$ on top of one another. $\lambda_{min}(A)$ denotes the minimum eigenvalue of A, and $trace(A)$ represents the trace of matrix A. $\mathbb{S}^n_+$ denotes the set of positive semi-definite matrices of order n.

II. Preliminaries

In this section, some background and preliminaries on behavioral systems theory [3] and safety in the context of CBCs [10] are provided. This sets the stage for the development of the proposed safety-certified data-driven MPC.

A. Behavioral Systems Theory

The following definition characterizes the minimal representation of LTI systems based on a sequence of input-state measurements.

Definition 1 [3]: Consider an LTI system G with a minimal realization represented by matrices $(A, B, C = 1, D = 0)$. An input-state sequence $\{u_k, x_k\}_{k=0}^{N-1}$ defines a trajectory of the system if there exists an initial state $\bar{x} \in \mathbb{R}^n$ satisfying the following conditions:

$$x_{k+1} = Ax_k + Bu_k, \quad x_0 = \bar{x}, \ k \in I_{[0,N-1]}. \tag{1}$$

Here, $x_k \in \mathbb{X} \subseteq \mathbb{R}^n$ denotes the state of the system, and $u_k \in \mathbb{U} \subseteq \mathbb{R}^m$ is the control input.

Given a sequence $x_{[0,N-1]}$, the associated Hankel matrix $H_L(x_{[0,N-1]})$ with the length of L is defined as [12]

$$H_L\big(x_{[0,N-1]}\big) := \begin{bmatrix} x_0 & x_1 & \cdots & x_{N-L} \\ x_1 & x_2 & \cdots & x_{N-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{L-1} & x_L & \cdots & x_{N-1} \end{bmatrix}. \tag{2}$$

Definition 2 [Persistence of Excitation (PE) Condition]: The sequence $u_{[0,N-1]} \in \mathbb{R}^{m\times N}$ is called as PE of order L if $rank(H_L(u_{[0,N-1]})) = mL$.

In this letter, we extensively apply the results of the following theorem derived from behavioral system theory [3], [9].

Theorem 1 [3]: Consider an LTI system G, for which a trajectory $\{u_k^d, x_k^d\}_{k=0}^{N-1}$ is given. Let $u^d_{[0,N-1]}$ be PE of order $L+n$. Under these conditions, a trajectory $\{\bar{u}_k, \bar{x}_k\}_{k=0}^{L-1}$ will be a valid trajectory of G if and only if there exists $\alpha \in \mathbb{R}^{N-L+1}$ such that

$$\begin{bmatrix} H_L\big(u^d_{[0,N-1]}\big) \\ H_L\big(x^d_{[0,N-1]}\big) \end{bmatrix} \alpha = \begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{x}_{[0,L-1]} \end{bmatrix}. \tag{3}$$

B. Safety and Control Barrier Certificates

Define the initial set $\mathbb{X}_0 \subset \mathbb{X}$ and the unsafe set $\mathbb{X}_u \subset \mathbb{X}$, where all are assumed to be bounded, and $\mathbb{X}_0 \cap \mathbb{X}_u = \emptyset$. Also, it is assumed that $\mathbf{0} \in \mathbb{X} \setminus \mathbb{X}_u$.

Definition 3 (Safety [4]): Given the system (1), for all initial states $x_0 \in \mathbb{X}_0$, the system (1) is called safe if there is no time horizon $T \in \mathbb{N}_0$ in which the trajectories starting from $\mathbb{X}_0$ reach the unsafe region $\mathbb{X}_u$.

To guarantee safety in autonomous systems, the notion of CBCs is given [5] as follows.

Definition 4 [2], [4]: The function $h : \mathbb{R}^n \to \mathbb{R}$ is called a CBC if the following properties hold:

$$h(x) \geq 0, \quad \forall x \in \mathbb{X}_0, \tag{4a}$$

$$h(x) < 0, \quad \forall x \in \mathbb{X}_u, \tag{4b}$$

$$\inf_u \{h(Ax + Bu)\} \geq (1 - \gamma)h(x), \quad \forall x \in \mathbb{X}, \tag{4c}$$

where $0 < \gamma \leq 1$.

Lemma 1 [4]: If $h(x)$ is a CBC, then the system (1) is safe in the sense of Definition 3.

Define the safe set $\mathbb{O}_\infty = \{x \in \mathbb{X} | h(x) \geq 0\}$ as the zero super-level set of $h(x)$. According to Lemma 1 and [2], if $h(x)$ is a CBC, one can easily conclude that the safe set $\mathbb{O}_\infty$ is a positively invariant set for the control system (1), i.e., $x_0 \in \mathbb{O}_\infty \Rightarrow x_k \in \mathbb{O}_\infty, \forall k \in \mathbb{N}_0$.

III. Problem Statement

A. Data-Driven MPC With CBCs

MPC, commonly used for obstacle avoidance, faces challenges in selecting the optimal prediction horizon. A short horizon may bring the system dangerously close to unsafe boundaries, denoted by $\mathbb{X}_u$, while a longer horizon increases computational complexity. Integrating MPC with CBCs can shorten the horizon [2], but requires precise knowledge of the system dynamics, often impractical in real settings. Moreover, this integration might lead to infeasibility [6]. This letter introduces a data-driven, safety-certified MPC using Theorem 1, replacing system dynamics with a data-driven constraint, ensuring both feasibility and stability, as follows

$$J_L^*(x_t) = \min_{\substack{\alpha(t),\ 0\leq\omega_k(t)\leq 1 \\ \bar{u}(t),\ \bar{x}(t)}} \sum_{k=0}^{L-1} \ell(\bar{u}_k, \bar{x}_k) + \psi(\omega_k(t)) + \bar{x}_L^T P \bar{x}_L, \tag{5a}$$

$$\text{s.t.} \begin{bmatrix} \bar{u}_{[0,L-1]}(t) \\ \bar{x}_{[0,L-1]}(t) \end{bmatrix} = \begin{bmatrix} H_L\big(u^d_{[0,N-1]}\big) \\ H_L\big(x^d_{[0,N-1]}\big) \end{bmatrix} \alpha(t), \tag{5b}$$

$$\bar{x}_0(t) = x_t, \ \bar{x}(t) \in \mathbb{X}, \ \bar{u}(t) \in \mathbb{U}, \tag{5c}$$

$$\bar{x}_L \in \mathbb{O}_\infty, \tag{5d}$$

$$h(\bar{x}_{k+1}(t)) \geq \omega_k(t)(1 - \gamma)h(\bar{x}_k(t)), \quad k \in I_{[0,L-1]}, \tag{5e}$$

where $\ell(\bar{u}, \bar{x}) = \|\bar{u}\|_R^2 + \|\bar{x}\|_Q^2$ with $Q, R > 0$, and $\psi(\omega_k)$ is defined as $\psi(\omega_k) = P_\omega(1 - \omega_k)^2$ with $P_\omega \geq 0$. $\bar{x}_L^T P \bar{x}_L$ is defined as the optimal unconstrained cost-to-go function and $\mathbb{O}_\infty$ denotes the safe terminal invariant set, which will be designed in Sections III-B and III-C, respectively. Compared to [3], it is assumed that the system's internal state is directly measurable. Even though disturbances and noises are not considered in (1), the presented method can be extended to consider measurement noise in the collected data samples u^d and x^d by following [3].

The inclusion of the CBC constraint (5e) without the adaptive parameter $\omega_k(t)$ has the potential to render the MPC problem (5) infeasible, as discussed in [2]. The inclusion of $\psi(\omega_k(t))$ in the cost function stems from a preference to maintain $\omega_k(t)$ at 1, thereby preserving the unchanged effect of the safety parameter γ. Nonetheless, in order to guarantee recursive feasibility—a point that will be proved in Section IV—it is imperative to appropriately adapt $\omega_k(t)$.

Assumption 1: The input $u^d_{[0,N-1]}$ is PE with an order of $L+n$ (See Definition 2).

Remark 1: The proposed safety-certified data-driven MPC incorporates three key techniques: the utilization of CBCs, the application of the behavioral system theorem, and the construction of a maximal safety-certified invariant set $\mathbb{O}_\infty$. CBCs effectively reduce computational complexity by shortening prediction horizons, particularly in obstacle avoidance scenarios [2]. Unlike the methods presented in [2], [6], our approach ensures recursive feasibility and exponential stability through the addition of specific terms $\omega_k(t)$, $\psi(\omega_k(t))$, and $\bar{x}_L^T P \bar{x}_L$, the proof of which will be provided in Section IV. Notably, the proposed method offers a fully data-driven approach for constructing $\mathbb{O}_\infty$ and obtaining the CBC $h(x)$, thereby making the proposed controller entirely data-driven.

B. Model-Free Optimal Control Problem Using Off-Policy RL

This subsection summarizes the key findings from [13], focusing on determining the optimal cost-to-go function $V(x) = x^T P x$ and the corresponding optimal control input $u^*(x) = -K^* x$, using off-policy RL for system (1).

The infinite-horizon performance function for system (1) is defined as $J(x_L, u_L) = \sum_{i=L}^{\infty} x_i^T Q x_i + u_i^T R u_i$.

Assumption 2: The pair (A, B) is stabilizable, and the pair $(A, \sqrt{Q})$ is detectable.

Under Assumption 2, a unique state-feedback optimal control policy $u^*(x) = -K^* x$ exists, where $K^* = (R + B^T P B)^{-1} B^T P A$. This policy minimizes the cost function J. $P > 0$ is the unique solution to the discrete algebraic Riccati equation [14]. By employing the optimal controller $u^*(x)$, the optimal performance becomes $V(x_L) = x_L^T P x_L$.

Since the system's dynamic is unknown, the following procedures are given using off-policy RL method [13] to find P and K^*, which is summarized in Algorithm 1.

First, $s \geq n^2 + m^2 + 2m + nm$ input-state data samples $\{u_k^p, x_k^p\}_{k=0}^{s-1}$ are collected offline by applying a stabilizing behavior policy $u^p = -K^0 x^p$. At each iteration j, the least squares (LS) method is employed to solve the following equation for P^{j+1}, L_2^{j+1}, and L_3^{j+1}.

$$\left[vec\left(P^{j+1}\right)^T \; vec\left(L_2^{j+1}\right)^T \; vec\left(L_3^{j+1}\right)^T \right]^T = ((\psi^j)^T \psi^j)^{-1} (\psi^j)^T \phi^j, \tag{6}$$

where ϕ^j and ψ^j are defined as follows

$$\phi^j = \begin{bmatrix} x_0^T Q x_0 + x_0^T (K^j)^T R K^j x_0 \\ x_1^T Q x_1 + x_1^T (K^j)^T R K^j x_1 \\ \vdots \\ x_{s-1}^T Q x_{s-1} + x_{s-1}^T (K^j)^T R K^j x_{s-1} \end{bmatrix},$$

$$\psi^j = \begin{bmatrix} E_{(xx)1} & E_{(xu)1} & E_{(uu)1} \\ E_{(xx)2} & E_{(xu)2} & E_{(uu)2} \\ \vdots & \vdots & \vdots \\ E_{(xx)s} & E_{(xu)s} & E_{(uu)s} \end{bmatrix}, \tag{7}$$

with $E_{(xx)i} = x_{i-1}^T \otimes x_{i-1}^T - x_i^T \otimes x_i^T$, $E_{(xu)i} = 2(x_{i-1}^T \otimes (u_{i-1} + K^j x_{i-1})^T)$, $E_{(uu)i} = (u_{i-1} - K^j x_{i-1})^T \otimes (u_{i-1} - K^j x_{i-1})$, where $i = 1, \ldots, s$.

Algorithm 1 Optimal Controller Design Using Off-Policy RL

1: Set the iteration number $j = 0$, and start with a stabilizing behavior policy $u^p = -K^0 x + e$, where e is a probing noise. Collect $s \geq n^2 + m^2 + 2m + nm$ samples (x_i, u_i), $i = 1, \ldots, s$.
2: Solve (6) by applying LS method for L_2^{j+1} and L_3^{j+1}.
3: Update the control policy K^{j+1} as (8).
4: Stop if $\|K^{j+1} - K^j\|_2 \leq \epsilon$, where $\epsilon > 0$. Otherwise, set $j = j+1$, and go back to step 2.

Assumption 3: The matrix ψ^j in (7) must have full column rank.

By finding L_2^{j+1} and L_3^{j+1}, one has:

$$K^{j+1} = \left(R + L_2^{j+1}\right)^{-1} L_3^{j+1}. \tag{8}$$

By iterating the above procedure, under Assumption 2, one has $K^* = K^j$ and $P = P^j$, as $j \to \infty$ [13].

Remark 2: The behavior policy u^p must satisfy three essential conditions. First, it should be stabilizable. To achieve this for unknown linear systems, a comprehensive method is outlined in [15]. Second, the policy should be rich enough to meet the requirements of Assumption 3. This can be accomplished by adding probing noise e to the behavior policy. Third, safety must be ensured during the data collection phase. To guarantee safety while collecting data samples, one can obtain the most permissive sub-level set of the Lyapunov function associated with the stabilizing policy [15] such that there is no intersection with unsafe sets. This set, being invariant, allows safe data collection when applying u_k^p, $k \in \mathbb{N}_0$.

C. Find $\mathbb{O}_\infty$ Using Control Barrier Certificates (CBCs)

In this subsection, the goal is to find a safety-certified maximal invariant set $\mathbb{O}_\infty = \{x \in \mathbb{X} | \; h(x) \geq 0\}$ given the data-driven unconstrained optimal control policy $u^*(x) = -K^* x$ obtained from Section III-B using CBCs. As a result, the following problem is given:

Problem 1: Given the closed-loop system (1) comprised of the unconstrained optimal control policy $u^*(x) = -K^* x$, find the most permissive CBC $h^*(x)$ satisfying (4), and as a result constructing the maximal invariant set $\mathbb{O}_\infty^* = \{x \in \mathbb{X} | \; h^*(x) \geq 0\}$.

To solve Problem 1, we assume that $h(x)$ is a polynomial. Therefore, according to [11], $h(x)$ can be represented as a square matrix $h(x) = Z(x)^T Q_h Z(x)$, where $Z(x)$ is a vector of monomials, and $Q_h \in \mathbb{R}^{k \times k}$ is a symmetric matrix. Also, the unsafe region $\mathbb{X}_u$ is described with multiple polynomials as

$$\mathbb{X}_u = \{x \in \mathbb{X} | q_i(x) < 0, \forall i \in \mathbb{M}\}, \quad \mathbb{M} = \{1, \ldots, M\}, \tag{9}$$

where M is the number of unsafe sets.

By using Definition 4, and considering the fact that $Vol(\mathbb{O}_\infty) \approx trace(Q_h)$ [11, Sec. 4.4.1], the following optimization problem is developed to solve Problem 1

$$h^*(x) = \underset{Q_h = Q_h^T}{\operatorname{argmax}} \; trace(Q_h), \tag{10a}$$

$$s.t. \; (4a) - (4b), \tag{10b}$$

$$C(x) = h(A_k x) + (\gamma - 1) h(x) \geq 0, \forall x \in \mathbb{X}, \tag{10c}$$

where $A_k = A - BK^*$. The maximal safe invariant set $\mathbb{O}^*_\infty$, obtained through (10), offers three advantages:

1) *System Safety:* The system remains safe within the confines of $\mathbb{O}^*_\infty$, as defined by Definition 3.
2) *Optimal Performance:* The system operates at an optimal level of performance due to applying optimal control input $u^*(x) = -K^*x$ in (10c).
3) *Larger Invariant Set Size:* As shown in [10], an invariant set derived through CBCs is at least as large as the sublevel set of $V(x)$.

However, there are two challenges in solving (10). First, solving the optimization problem (10) in general is difficult since checking non-negativity is often considered as a non-trivial problem [16]. Second, the system's dynamic is required in (10c), which is not available in this letter. The initial identification of system dynamics A and B for evaluating (10c) has two limitations: (1) Errors in identification impact the calculation of the invariant set $\mathbb{O}_\infty$, requiring robustness analysis. (2) Identifying the A and B, which contain $n^2 + nm$ unknowns, demands more data than the scalar $C(x)$.

To solve the first challenge, by applying P-Satz Lemma [10, Lemma 3.3], (10) can be solved by the following optimization problem

$$\min_{\substack{Q_h=Q_h^T,\ L(x)\in\mathbb{P}^{SOS}\\ J_j(x)\in\mathbb{P}^{SOS},\ \forall j\in\mathbb{M}}} -\, trace(Q_h),$$
$$\text{s.t. } h(A_kx) + (\gamma - 1)h(x) - L(x)h(x) \in \mathbb{P}^{SOS}, \quad (11a)$$
$$-\, h(x) + J_j(x)q_j(x) \in \mathbb{P}^{SOS}. \quad (11b)$$

In (11), we assume that $\mathbb{X} = \mathbb{R}^n$ and $\mathbb{X}_0 = \mathbb{O}_\infty$, given that $\mathbb{X}_0 \subseteq \mathbb{O}_\infty$. To guarantee that the condition in (10c) is satisfied over a bounded set, $\mathbb{O}_\infty$ is considered as that bounded set.

To address the second challenge associated with solving (10)—that of requiring the system dynamics—one can reformulate (10c) as follows:

$$C(x_k) = h(x_{k+1}) + (\gamma - 1)h(x_k) \geq 0, \quad k \in \mathbb{N}_0 \quad (12)$$

where $x_{k+1} = (A - BK^*)x_k$. Since $C(.)$ is dependent on the system dynamics, and that is unknown, $C(.)$ is approximated as $C(x) = {I_C}^T\vec{m}_{0,d}(x)$, where $I_C \in \mathbb{R}^{n_1}$ is an unknown vector and $\vec{m}_{0,d}(x)$, where $x \in \mathbb{R}^n$, represents a vector of all distinct monic polynomials, arranged in lexicographic order, with degrees ranging from 0 to d. As a result, one has

$$\big(Z(x_{k+1})^T \otimes Z(x_{k+1})^T + (\gamma - 1)Z(x_k)^T \otimes Z(x_k)^T\big)vec(Q_h) = \vec{m}_{0,d}(x_k)^T I_C. \quad (13)$$

In (13), there are n_1 unknown parameters. As a result, a minimum of n_1 data samples is needed to determine I_C. By collecting $N_s \geq n_1$ data samples through applying the control policy $u^*(x)$, (13) can be rewritten as:

$$\Omega_S I_C = Y_S(Q_h), \quad (14)$$

where $Y_S = \delta I_{ZZ} vec(Q_h)$. δI_{ZZ} is defined as $\delta I_{ZZ} = \big[\psi_0 \ldots \psi_{N_s-1}\big]^T$ with $\psi_i = (Z(x_{i+1}) \otimes Z(x_{i+1})) + (\gamma - 1)Z(x_i) \otimes Z(x_i)$, and Ω_S is defined as $\Omega_S = \big[\vec{m}_{0,d}(x_i)^T, \ldots, \vec{m}_{0,d}(x_i)^T\big]^T$, $i = 0, \ldots, N_s - 1$.

Assumption 4: The data samples are collected such that Ω_S has full column ranks.

Algorithm 2 Solving the Bilinear SOS Program (15)

1: **Input:** Unsafe region $\mathbb{X}_u$, $Z(x)$ as a vector of monomials, $0 < \gamma \leq 1$, P, and $u^*(x) = -K^*x$.
2: **Output:** The most permissive safety barrier certificate $h^*(x) = Z(x)^T Q_h^* Z(x)$ and $\mathbb{O}^*_\infty$.
3: **Collect offline data samples:** Apply $u^*(x) = -K^*x$ to system (1) and collect N_s data samples satisfying Assumption 4. Then, find Ω_S and δI_{ZZ}.
4: **Initialization:** Initialize $h(x)$, which could be $h^0(x) = c - x^T Px = Z(x)^T Q_h^0 Z(x)$, where $c > 0$ can be obtained as [10].
5: **while** *trace*(Q_h) is increasing **do**
6: Fix $h(x)$ and search for $L(x)$ by expanding the feasible space using the variable $\varepsilon \geq 0$ and solving the following SOS programming
$$\min_{\varepsilon\in\mathbb{R}^+,\ L(x)\in\mathbb{P}^{SOS}} -\, \varepsilon \quad (16a)$$
$$s.t.\ \Omega_S I_C = Y_S(Q_h), \quad (16b)$$
$${I_C}^T\vec{m}_{0,d}(x) - L(x)h(x) - \varepsilon \in \mathbb{P}^{SOS}, \quad (16c)$$
7: Fix $L(x)$ and search for $h(x)$ by solving (15).
8: **end while**

As a result, by substituting (11a) with (14), one can rewrite (11) with the following data-driven optimization problem

$$\max_{\substack{Q_h=Q_h^T,I_C\in\mathbb{R}^{n_1}\\ L(x),\ J_j(x)\in\mathbb{P}^{SOS},\ \forall j\in\mathbb{M}}} trace(Q_h),$$
$$\text{s.t. } \Omega_S I_C = Y_S(Q_h), \quad (15a)$$
$${I_C}^T\vec{m}_{0,d}(x) - L(x)h(x) \in \mathbb{P}^{SOS}, \quad (15b)$$
$$-\, h(x) + J_j(x)q_j(x) \in \mathbb{P}^{SOS}. \quad (15c)$$

The optimization problem (15) contains bilinear decision variables. Hence, Algorithm 2 is given to solve (15) efficiently using iterative search algorithm [10].

D. Design Procedure and Computational Complexity

The following offline procedure is taken to obtain P and $\mathbb{O}_\infty$, which will be used in the online implementation of (5).

Offline Procedure: In this letter, first, $s \geq n^2+m^2+2m+nm$ data samples are collected satisfying Assumption 3 to find the unconstrained optimal cost-to-go function $V^*(x) = x^T Px$ and control policy $u^*(x)$ (See Section III-B for more details). Subsequently, the obtained $u^*(x)$ is applied to the unknown system described by (1). This results in the collection of $N_s \geq n_1$ data samples, which satisfy Assumption 4. From these data samples, we construct the data-driven safety-certified maximal invariant set $\mathbb{O}_\infty$ (See Algorithm 2). For improved data efficiency, one can select N input-state data samples from the collected data sets, provided Assumption 1 is met.

Computational Complexity: This letter contains two offline Algorithms 1-2 to determine P and $\mathbb{O}_\infty$, and one online MPC implementation (5).

The online data-driven MPC (5) consists of $N_{MPC} = (m + n)L + (n + N) + 1$ decision variables. The computational complexity of solving this generic nonlinear programming (5) by using sequential quadratic programming (SQP) [17] can be approximated by expected floating point operations per second (flops). In the worst case, the number of flops required to solve each QP subproblem of (5) is approximated by $flop_{QP} = i_{IP} \times (\frac{2}{3}N_{MPC}^3 + 2N_{MPC}^2)$, where i_{IP} is the number of interior point iterations and is expected to be $\mathcal{O}(\sqrt{N_{MPC}}\log(\frac{1}{\epsilon}))$ for the ϵ-accurate solution [18].

The computational cost of off-policy RL given in Algorithm 1, which is mainly dependent on performing the LS method on (6), is obtained as $\mathcal{O}((\frac{(n+m)(n+m+1)}{2})^2 s)$ [19]. Algorithm 2 consists of solving a bilinear SOS program (15) which has $\frac{k(k+1)}{2} + n_1 + \frac{n_L(n_L+1)}{2} + \frac{n_J(n_J+1)}{2}$ decision variables, where we assume $L(x) = L_b(x)^T L^Q L_b(x)$ and $J(x) = J_b(x)^T J^Q J_b(x)$, with $L^Q \in \mathbb{S}_+^{n_L}$ and $J^Q \in \mathbb{S}_+^{n_J}$. While SOS programming faces scalability issues, the diagonally dominant-sum-of-squares technique proposed by [20] converts the optimization problem into linear programming, enhancing efficiency for high-dimensional systems. According to [7], reducing online computational cost is achieved by expanding the terminal set, which shortens the time for the system to reach it. In this letter, we achieve the expansion of the terminal invariant set through the concept of CBCs, thus reducing the online computational burden.

IV. Closed-Loop Theoretical Properties

In this section, we investigate the recursive feasibility and exponential stability of the closed-loop system (1).

Theorem 2 (Recursive Feasibility): Under Assumptions 1-4, and assuming that $u^*(x) \in \mathbb{U}$, $\forall x \in \mathbb{O}_\infty$, if the proposed safety-certified data-driven MPC (5) is feasible at initial time $t = 0$, then it is feasible at any $t \in \mathbb{N}_0$.

Proof: Consider the sequence of feasible states and inputs at time instant $t = 0$ defined as

$$\boldsymbol{x}_s^0 = [\bar{x}_0^0 \quad \bar{x}_1^0 \quad \dots \quad \bar{x}_L^0],\ \boldsymbol{u}_s^0 = [\bar{u}_0^0 \quad \bar{u}_1^0 \quad \dots \quad \bar{u}_{L-1}^0]$$

By appending $\bar{x}_{L+1}^0$ and $\bar{u}_L^0(x) = -K^*x$ into $\boldsymbol{x}_s^0$ and $\boldsymbol{u}_s^0$, the shifted sequence of states and inputs are defined as

$$\boldsymbol{x}_s^{0+} = [\bar{x}_1^0 \quad \bar{x}_2^0 \quad \dots \quad \bar{x}_{L+1}^0],\ \boldsymbol{u}_s^{0+} = [\bar{u}_1^0 \quad \bar{u}_2^0 \quad \dots \quad \bar{u}_L^0]$$

We need to show that $\boldsymbol{x}_s^{0+}$ and $\boldsymbol{u}_s^{0+}$ are feasible solutions of the proposed MPC (5) at time step $t = 1$. Given that $\mathbb{O}_\infty$ is an invariant set under the definition of CBC (Definition 4) for the closed-loop optimal control system G, it follows that $\bar{x}_{L+1}^0 \in \mathbb{O}_\infty$, which satisfies (5d). Additionally, under the theorem assumptions, $\bar{u}_L^0(\bar{x}_L^0) = -K^*\bar{x}_L^0 \in \mathbb{U}$, satisfying (5c). For (5b), Theorem 1 and Assumption 1 ensure the existence of $\alpha(1)$ that satisfies equality (5b). Regarding (5e), the condition $h(\bar{x}_L^0) \geq 0$ follows from Definition 4, and together with the term $\omega_{L-1}(1)(1-\gamma) \geq 0$, implies $h(\bar{x}_{L+1}^0) \geq 0$. This can be achieved by setting $\omega_{L-1}(1) = 0$, which justifies the adaptive change of this parameter to ensure recursive feasibility when no other solutions are available. Therefore, we have shown that if the proposed MPC (5) is feasible at $t = 0$, it remains feasible at $t = 1$. By following the same procedures, one can easily show that the proposed MPC (5) is recursively feasible, i.e., feasibility at time step $t = 0$ results in feasibility at any time instant $t \in \mathbb{N}$. Assumptions 1-4 are necessary for the existence of K^*, $\mathbb{O}_\infty$, and $\alpha(t)$. ∎

Assumption 5: There exists some constant $\Gamma > 0$ such that $J_L^*(x) \leq \Gamma$, $\forall x \in \mathbb{X}_L$, where $\mathbb{X}_L = \{x \in \mathbb{X} |\ J_L^*(x) \leq \infty\}$.

Remark 3: Compared to [3], where there must exist $c_u > 0$ such that $J_L^*(x) \leq c_u\|x\|_2^2$, Assumption 5 is less restrictive.

Theorem 3 (Exponential Stability): Under Assumptions 1-5, the closed-loop system G under the proposed MPC (5) is locally exponentially stable.

Proof: Consider the Lyapunov function candidate $V^*(x) = J_L^*(x)$. By using (5a), one has

$$V^*(x_t) = \ell(u_t^*, x_t) + \psi(\omega_0^*(t)) + \sum_{k=1}^{L-1} \ell(\bar{u}_k^*, \bar{x}_k^*) + \psi(\omega_k^*(t)) + (\bar{x}_L^*)^T P \bar{x}_L^*$$
$$\geq \ell(u_t^*, x_t) + \sum_{k=1}^{L-1} \ell(\bar{u}_k^*, \bar{x}_k^*) + \psi(\omega_k^*(t)) + (\bar{x}_L^*)^T Q \bar{x}_L^*$$
$$+ (\bar{x}_L^*)^T \left(K^{*T} R K^*\right) \bar{x}_L^* - Z_v \pm (\bar{x}_L^*)^T (A - BK^*)^T P (A - BK^*) \bar{x}_L^*$$

where $Z_v = (\bar{x}_L^*)^T (Q - P + F^T R F) \bar{x}_L^*$. Since $Z_v + (\bar{x}_L^*)^T (A - BK^*)^T P (A - BK^*) \bar{x}_L^* \leq 0$ according to [13], the above equation can be rewritten as

$$V^*(x_t) \geq \ell(u_t^*, x_t) + V^*(x_{t+1}) - \psi(\omega_L^*(t)). \qquad (17)$$

Since $\bar{x}_L^* \in \mathbb{O}_\infty$, one has $\omega_L^* = 1$. As a result, $\psi(\omega_L^*(t)) = 0$. Thus, (17) can be simplified as

$$V^*(x_{t+1}) - V^*(x_t) \leq -\lambda_{min}(Q) x_t^T x_t. \qquad (18)$$

$V^*(x)$ is lower bounded by $\lambda_{min}(Q)x^T x$, i.e., $V^*(x) \geq \lambda_{min}(Q)x^T x$. The following procedures are given to find the upper bound of $V^*(x)$. By applying the results from Theorem 2, $u_t(x) = -K^* x_t$ is a feasible control policy for the data-driven MPC (5). As a result, one has

$$V^*(x) \leq \sum_{k=0}^{L-1} \ell(-K^*\bar{x}_k, \bar{x}_k) + \bar{x}_L^T P \bar{x}_L. \qquad (19)$$

By assuming $x \in \mathbb{O}_\infty$, one has $\psi(.) = 0$ in the above inequality. Also, with a slight abuse of notation, in (19), we denote $\bar{x}_k$ as the closed-loop state of the system resulting from applying the unconstrained optimal control policy $-K^*\bar{x}_k$. By using the property of the invariant set $\mathbb{O}_\infty$, (19) can be rewritten as

$$V^*(x) \leq (L\theta + \|P\|_2)\|x_0\|_2^2, \quad \forall x_0 \in \mathbb{O}_\infty, \qquad (20)$$

where $\theta = \|Q + (K^*)^T R K^*\|$. To derive an upper bound for $V^*(x)$ for all $x_0 \in \mathbb{X}_L$, we recall the results given in the proof of [21, Th. 7.4] along with applying Assumption 5. Hence, one has

$$V^*(x) \leq \left(\frac{\Gamma\beta}{\hat{\beta}}\right)\|x_0\|_2^2, \quad \forall x_0 \in \mathbb{X}_L, \qquad (21)$$

where $\hat{\beta} = \max\{\beta\|x\|_2^2 |\ \|x\|_2 \leq d\}$, and $d > 0$ is defined as $d \in \{d_1 > 0 | \{\|x\|_2 \leq d_1\} \subset \mathbb{O}_\infty,\ x \in \mathbb{R}^{n_x}\}$. By combining the results given in (18), (21), and the obtained lower-bound for $V^*(x)$, i.e., $V^*(x) \geq \lambda_{min}(Q)x^T x$, one can conclude from the standard Lyapunov arguments that the closed-loop system G under the proposed MPC controller (5) is locally exponentially stable. ∎

V. Simulation Results

Consider a two-dimensional single integrator with a sampling time $\Delta T = 0.2$. The unsafe set $\mathbb{X}_u$ is defined as $\mathbb{X}_u = \{(x_1, x_2) : (x_1 - 0.25)^2 + (x_2 - 0.45)^2 - 0.1^2 < 0\}$. The performance parameters are $Q = I_{2\times2}$, $R = 10I_{2\times2}$, and $P_\omega = 100$. The safe set kernel is $Z(x_1, x_2) = [1\ x_1\ x_2]^T$. The initial condition is $x_0 = [0.8, 1.2]^T$, and $L = 6$.

Offline Data Collection: We conducted data collection in two stages. Initially, we gathered $s = 20$ data samples,

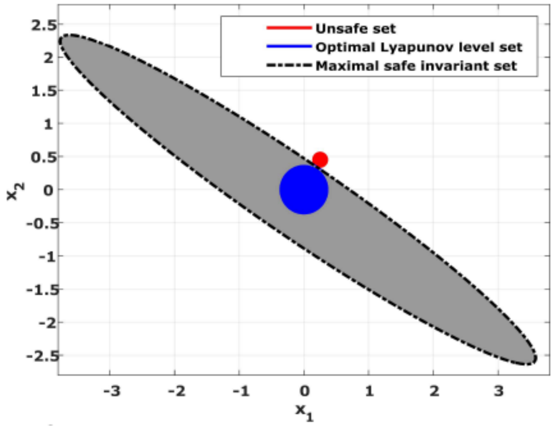

Fig. 1. Comparing the size of maximal safe invariant set $\mathbb{O}^*_\infty$ with the Lyapunov sublevel set Ω_c.

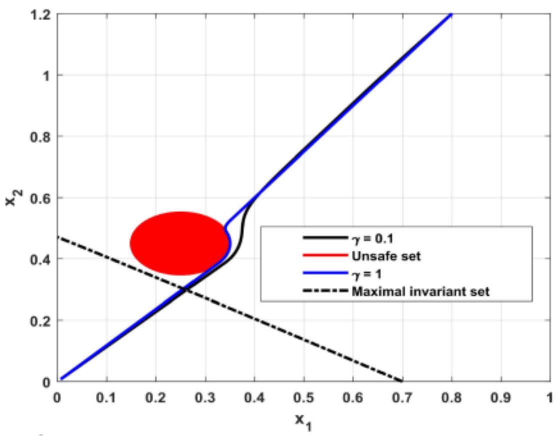

Fig. 2. Comparison of the trajectories by changing γ.

ensuring $s \geq n^2 + m^2 + 2m + nm$ from the given system using the stabilizing behavior policy $u_k^p = -K^0 x_k^p + e_k$, with $K^0 = 0.2I_{2\times 2}$. The probing noise, e_k, meets Assumption 3. Using Algorithm 1, we determined $K^* = 0.31I_{2\times 2}$ and $P = 16.32I_{2\times 2}$. Once the optimal control policy was derived, we identified the maximal safe invariant set $\mathbb{O}^*_\infty(x)$, given by $h^*(x)$. Algorithm 2 was applied to collect $N_s = 20$ data samples using the optimal control input satisfying Assumption 4. Then, $N = 25$ input-state measurements are collected from the offline data sets satisfying Assumption 1.

In Fig. 1, the maximal safe invariant set $\mathbb{O}^*_\infty$ is depicted as a grey-filled dashed black curve, while the Lyapunov sublevel set $\Omega_c = \{x \in \mathbb{R}^2 | x^T P x \leq c\}$ with $c = 2.5$ is shown as a blue ellipse. The unsafe set is represented by a red ellipse. The figure illustrates that the volume of $\mathbb{O}^*_\infty$ is considerably larger than the Lyapunov sublevel set, leading to a larger feasible set in the proposed MPC controller (5). Fig. 2 shows how the trajectories deviate from the unsafe set by changing γ. Fig. 2 clearly shows that by increasing γ, the trajectory is getting close to the unsafe set, which results in a potentially unsafe action. Reducing γ enhances safety by keeping trajectories away from unsafe sets, though it may compromise optimal performance.

VI. Conclusion

This letter presents a safety-certified data-driven MPC for LTI systems. A safe maximal terminal invariant set is constructed by using CBCs, resulting in decreasing the online computational cost. Thanks to applying behavioral system theory, the requirement of knowing the exact system dynamics is relaxed. Proofs concerning the stability and feasibility are proposed, and the effectiveness of the proposed method is demonstrated through a simulation. Future work may include enhancing robustness against measurement noise and adapting the method to linear time-varying systems.

References

[1] B. Kouvaritakis and M. Cannon, *Model Predictive Control*, vol. 38. Cham, Switzerland: Springer Int. Publ., 2016.

[2] J. Zeng, B. Zhang, and K. Sreenath, "Safety-critical model predictive control with discrete-time control barrier function," in *Proc. Am. Control Conf. (ACC)*, 2021, pp. 3882–3889.

[3] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Trans. Autom. Control*, vol. 66, no. 4, pp. 1702–1717, Apr. 2020.

[4] S. Prajna, A. Jadbabaie, and G. J. Pappas, "A framework for worst-case and stochastic safety verification using barrier certificates," *IEEE Trans. Autom. Control*, vol. 52, no. 8, pp. 1415–1428, Aug. 2007.

[5] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *Proc. 18th Eur. control Conf. (ECC)*, 2019, pp. 3420–3431.

[6] J. Zeng, Z. Li, and K. Sreenath, "Enhancing feasibility and safety of nonlinear model predictive control with discrete-time control barrier functions," in *Proc. 60th IEEE Conf. Decis. Control (CDC)*, 2021, pp. 6137–6144.

[7] D. Limon, T. Alamo, and E. F. Camacho, "Enlarging the domain of attraction of MPC controllers," *Automatica*, vol. 41, no. 4, pp. 629–635, 2005.

[8] P. Tooranjipour and B. Kiumarsi, "Constructing safety barrier certificates for unknown linear optimal control systems," in *Proc. IEEE 17th Int. Conf. Control Autom. (ICCA)*, 2022, pp. 213–219.

[9] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Syst. Control Lett.*, vol. 54, no. 4, pp. 325–329, Apr. 2005.

[10] L. Wang, D. Han, and M. Egerstedt, "Permissive barrier certificates for safe stabilization using sum-of-squares," in *Proc. Annu. Am. Control Conf. (ACC)*, 2018, pp. 585–590.

[11] G. Chesi, *Domain of Attraction: Analysis and Control via SOS Programming*, vol. 415. London, U.K.: Springer, 2011.

[12] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems: Theory—Implementation—Applications*. New York, NY, USA: Springer, 2012.

[13] B. Kiumarsi, F. L. Lewis, and Z.-P. Jiang, "H∞ control of linear discrete-time systems: Off-policy reinforcement learning," *Automatica*, vol. 78, pp. 144–152, Apr. 2017.

[14] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*. Hoboken, NJ, USA: Wiley, 2012.

[15] A. Lamperski, "Computing stabilizing linear controllers via policy iteration," in *Proc. 59th IEEE Conf. Decis. Control (CDC)*, 2020, pp. 1902–1907.

[16] G. Blekherman, P. A. Parrilo, and R. R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*. Philadelphia, PA, USA: SIAM, 2012.

[17] M. Diehl, H. J. Ferreau, and N. Haverbeke, "Efficient numerical methods for nonlinear MPC and moving horizon estimation," in *Nonlinear Model Predictive Control: Towards New Challenging Applications*. Berlin, Germany: Springer, 2009, pp. 391–417.

[18] C. Shen and Y. Shi, "Distributed implementation of nonlinear model predictive control for AUV trajectory tracking," *Automatica*, vol. 115, May 2020, Art. no. 108863.

[19] F. A. Yaghmaie, F. Gustafsson, and L. Ljung, "Linear quadratic control using model-free reinforcement learning," *IEEE Trans. Autom. Control*, vol. 68, no. 2, pp. 737–752, Feb. 2023.

[20] A. A. Ahmadi and A. Majumdar, "DSOS and SDSOS optimization: More tractable alternatives to sum of squares and semidefinite optimization," *SIAM J. Appl. Algebra Geom.*, vol. 3, no. 2, pp. 193–230, 2019.

[21] S. Singh, Y. Chow, A. Majumdar, and M. Pavone, "A framework for time-consistent, risk-sensitive model predictive control: Theory and algorithms," *IEEE Trans. Autom. Control*, vol. 64, no. 7, pp. 2905–2912, Jul. 2019.