Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-28 (2024) 402–407

Conflict-Aware Data-Driven Safe Linear Quadratic Control ⋆

Ramin Esmzad [*] **Nariman Niknejad** [*] **Hamidreza Modares** [*]

[*] *Department of Mechanical Engineering, Michigan State University, East Lansing, MI 48824 (e-mails: {esmzadra, niknejad,modaresh}@msu.edu).*

Abstract: This paper presents a new data-driven conflict-aware safe Linear Quadratic Regulator (LQR) with dual safety measures. During design, the LQR control gain is optimized solely from data to minimize costs and enlarge a conflict-free zone ensuring safe optimal trajectories. In execution, a control barrier certificate (CBC) verifies the safety of controller actions. The design-time intervention implicitly aligns LQR weights with safety constraints, preventing harmful conflicts and reducing the need for frequent CBC interventions. To achieve this, the LQR gain is parameterized with a λ-contractive safe set. Simulation results on the vehicle steering model demonstrate the effectiveness of this approach.

Keywords: LQR, Safe Control, $\lambda-$Contractive Control, Data-driven Control, Robustness, Optimal Control.

1. INTRODUCTION

Data-driven control represents a transformative shift from traditional methodologies, leveraging system data to inform control strategies amidst complex, uncertain, or variable system dynamics (Hou and Wang, 2013; Markovsky and Dörfler, 2021). Unlike conventional methods relying solely on mathematical models, data-driven approaches directly incorporate data to tailor control strategies for specific performance objectives.

The data-driven Linear Quadratic Regulator (LQR) emerges as a cornerstone in this methodology. Both indirect and direct data-based (Krishnan and Pasqualetti, 2021; De Persis and Tesi, 2021) methodologies are presented for learning the LQR solutions. Indirect data-driven LQR utilizes historical data for system identification, guiding subsequent LQR gain design through model-based methods. In contrast, direct data-driven LQR, known as model-free LQR, circumvents system identification, directly optimizing gain to minimize a cost function.

Many control systems are safety-critical systems for which the safety guarantee of the learned controller is crucial for their successful deployment on real-world applications (Lee, 2023). Ensuring that system states will remain within predefined safe regions and avoiding unsafe states or actions despite uncertainties requires robustness analysis, uncertainty quantification, and the integration of safety constraints.

Data-driven Safe optimal control, specifically safe LQR, has also been considered in the literature (Zanon and Gros, 2020; Choi et al., 2020). These approaches typically leverage add-on methods such as control-barrier certificates (CBC) methods (Ames et al., 2019; Agrawal and Sreenath, 2017) to certify the safety of an unconstrained optimal controller (e.g., the LQR controller) by correcting its actions in real-time. If the conflict between safety constraints and the performance function to be optimized is significant, this approach can lead to undesired behaviors, such as performance deterioration, infeasibility, or even unsafe actions. Designing conflict-aware LQR controllers can avoid frequent intervention requirements of CBCs, which is key to the success of future learning-enabled systems with safety constraints.

This paper introduces a novel data-driven conflict-aware safe LQR featuring dual safety interventions. At a higher level (or during design), the LQR control gain is solely data-driven, optimizing cost while enlarging a conflict-free zone to ensure safe trajectories. At a lower level (or in real-time), a CBC intervenes to certify the safety of controller actions. The higher-level intervention implicitly aligns LQR weights with safety constraints, mitigating destructive conflicts that would otherwise necessitate frequent CBC interventions.

Moreover, the paper presents a new data-based approach for learning the LQR control while optimizing the size of the conflict-free zone. This method entails a closed-loop representation of system dynamics using data and a decision variable, ensuring robust stability and a trade-off between optimality and safe zone size. The quality of collected data significantly impacts the size of the conflict-free zone and performance, as validated through simulation results on a vehicle steering model.

Additionally, λ-contractivity (Blanchini and Miani, 2008) emerges as a fundamental principle in designing safe controllers, ensuring deviations from nominal behavior induced by uncertainty or disturbances remain bounded. Integrating λ-contractive properties into data-driven LQR

⋆ 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.

10.1016/j.ifacol.2025.01.079

controllers mitigates instability risks, ensuring robust performance under uncertainty.

The paper's primary contribution lies in integrating safety into the LQR controller structure through the λ-contractivity principle. Parameterizing the LQR gain with a λ-contractive safe set and optimizing its volume determines the maximum safe optimal region within a given polyhedral safe set. Both model-based and direct data-driven safe LQR controllers are formulated in terms of Linear Matrix Inequalities (LMIs), with the maximum safe optimal set size heavily influenced by the quality of collected data.

This paper is organized as follows. Section 2 defines the safe LQR problems. Sections 3 and 4 provide the model-based and data-driven safe LQR formalism, respectively. In Section 5, a simulation example is conducted to show the superiority of the proposed approach. Finally, section 6 concludes the paper.

2. PROBLEM FORMULATION

Consider a time-invariant stochastic linear discrete-time system of the form

$$x_{k+1} = A\,x_k + B\,u_k + w_k, \tag{1}$$

where $k \in \mathbb{N}$, $x_k \in \mathbb{R}^n$ is the system's state, and $u_k \in \mathbb{R}^m$ is the control input. Moreover, A and B are unknown transition and input matrices of appropriate dimensions respectively. Furthermore, $w_k \in \mathbb{R}^n$ represents the system noise.

Assumption 1. The system noise ω_k is governed by the Gaussian distribution $\mathcal{N}(0, W)$ where $W \succ 0 \in \mathbb{R}^{n\times n}$ is its covariance matrix.

Assumption 2. The pair (A, B) is unknown but stabilizable.

The goal is to design a linear controller that stabilizes the system while respecting safety constraints despite uncertainties in A and B matrices. Since stability and safety are bare minimum requirements, optimal performance of the controller is typically achieved by optimizing a performance or cost function. Before formalizing the safe optimal control problem, the controller form, the cost function, and the safe set are defined next.

In this paper, the control input has the following form

$$u_k = K\,x_k, \tag{2}$$

where $K \in \mathcal{K}$ is the control gain and the set of stabilizing control gains is defined as $\mathcal{K} = \{K \in \mathbb{R}^{m\times n} : \rho(A + BK) < 1\}$.

The cost function that is used to assess the system performance is given by

$$J(K, x_0) = \sum_{k=0}^{\infty} \gamma^k \mathbb{E}\left[x_k^T Q x_k + u_k^T R u_k\right] \tag{3}$$

where $Q \in \mathbb{R}^{n\times n} \succ 0, R \in \mathbb{R}^{m\times m} \succ 0$. Also, $\gamma \in (0, 1)$ is the discount factor.

Assumption 3. For the system (1) and the cost function (3), the pair $(A, \sqrt{Q})$ is detectable.

Finally, the safe set is assumed to be a polyhedral set $\mathcal{P}$ defined as

$$\mathcal{P} = \left\{x \in \mathcal{R}^n \quad | \quad a_s^T x \leq 1, \quad s = 1, \cdots, q\right\}, \tag{4}$$

specifies the safe operational region of states.

Now, let us consider the following problems.

Problem 1. Consider the system (1). Design a controller (2) that optimizes the cost function (3) while respecting the safety constraints (4).

A significant challenge is that designing a feedback controller that solves Problem 1 is computationally intractable. Therefore, it is a common practice to learn a linear controller in the form of (2) without accounting for the safety constraints (i.e., to solve a standard LQR problem) and use CBCs (Ames et al., 2019) to myopically intervene with the actions of the LQR controller to correct their unsafe actions. That is, given u_k^* as the current action of the LQR controller at time k, CBC myopically intervenes with LQR by solving the following minimally-intervened safety-certified optimization

$$u_k^{so} = \arg\min_u (u - u_k^*)^T(u - u_k^*) \tag{5a}$$

$$\text{s.t.} \quad h(x_{k+1}) - (1-\eta)h(x_k) \geq 0. \tag{5b}$$

where u_k^{so} is the safe optimal control signal to be applied, $0 < \eta \leq 1$, and $h(x_k) \geq 0$ represents the safety region. We can have multiple $h_s(x_k) = 1 - a_s^T x_k \geq 0, s = 1, \cdots, q$ for the safe set in (4).

Unfortunately, when the conflict between the performance function and the safety constraints is huge, the low-level intervention step must be performed frequently, which can ruin the performance or even cause infeasibility or unsafe actions under uncertainties. To resolve this issue, which is a significant impediment in the realization of safe leaning-enabled controllers, a new safe-optimal control design is presented that learns a control gain that not only optimizes the LQR cost but also leads to a reasonable conflict with safety requirements, which can be resolved using CBC without the need for frequent intervention. To this end, the following problem is formalized. The following definition is required.

Definition 1. $\lambda-$Contractive and Positive Invariant Sets Consider the system (1). Let $\lambda \in (0, 1]$. The set $\mathcal{P}$ is considered λ-contractive if for any $x_k \in \mathcal{P}$ it holds that $x_{k+1} \in \lambda\mathcal{P}$. When $\lambda = 1$, $\mathcal{P}$ is positive invariant (Blanchini and Miani, 2008).

Problem 2. Consider the system (1). Suppose we have collected the following N sequences of data

$$U_0 = [u_0\ u_1\ \cdots\ u_{N-1}], \tag{6a}$$

$$X_0 = [x_0\ x_1\ \cdots\ x_{N-1}], \tag{6b}$$

$$X_1 = [x_1\ x_2\ \cdots\ x_N], \tag{6c}$$

where $U_0 \in \mathbb{R}^{m\times N}$, X_0 and $X_1 \in \mathbb{R}^{n\times N}$. The corresponding noise sequences $\Omega_0 \in \mathbb{R}^{n\times N}$ are

$$\Omega_0 = [\omega_0\ \omega_1\ \cdots\ \omega_{N-1}], \tag{7}$$

in which we have no access to it.

- *The Design time Phase:* Given the available data, learn a control gain K for (1) to optimize a weighted sum of the cost function (3) and the size of the largest $\lambda-$contractive ellipsoidal set $x_k^T P_s^{-1} x_k \leq 1$ contained in the safe set (4) (i.e, a conflict-free zone). This ensures performance guarantees over a conflict-free zone of reasonable size. We call this a conflict-aware controller.

- *The Real-time Phase:* Solve the optimization (5) to correct possible unsafe actions of the learned conflict-aware controller.

2.1 LQR Formulations Without Safety Considerations

We first discuss existing semi-definite program (SDP) based formulations for solving the LQR problem with no safety constraints. There are two main approaches to formulating the infinite-horizon stochastic LQR problem (3): dynamic programming (Ricatti equation) and state covariance approaches. The state covariance formulation can also be interpreted as the $\mathcal{H}_2-$norm minimization problem of the cost function (3). From the optimization theory perspective, the first approach is called the primal problem and the second one is the dual problem. In both formulations, K is the optimal LQR gain that minimizes the quadratic cost (3). The primal and dual SDP optimizations for solving the LQR problem are provided next.

LQR Primal Problem

$$\max_{K,Y} \quad Tr(W^{-1}Y) \tag{8a}$$

$$\text{s.t.} \quad \gamma(A+BK)^T P(A+BK) - P + Q + K^T RK \preceq 0. \tag{8b}$$

$$Y \succeq 0, \tag{8c}$$

In this approach, $Y = P^{-1}$, and P is the solution to the Ricatti equation.

LQR Dual Problem

$$\min_{\beta,K,W_c\succeq W,L} \quad \beta \tag{9a}$$

$$\text{s.t.} \quad \gamma(A+BK)W_c(A+BK)^T - W_c + W \preceq 0, \tag{9b}$$

$$L - KW_cK^T \succeq 0, \tag{9c}$$

$$tr(QW_c) + tr(RL) \leq \beta. \tag{9d}$$

In this formulation, W_c represents the state covariance matrix (or controllability Gramian).

In the subsequent sections, we first provide solutions to the model-based safe LQR that solves Problem 2 and then extend the results to the data-driven case. Only the design-time step (step 1 in Problem 2) is developed as the real-time step is similar to the existing CBC methods.

3. MODEL-BASED SAFE LQR

This section investigates different optimization problems to design a safe optimal controller for the case where we have a reliable knowledge of the system matrices, hence model-based safe LQR. We have categorized safe LQR designs into direct and indirect approaches. In the direct approach, we directly add the safety constraints to the LQR optimization problem. At the same time, in the indirect method, we parameterize the control gain K to satisfy safety constraints in addition to optimality concerns. So, we embed safety into the controller structure.

3.1 Direct Safe LQR

To directly solve Problem 1 using a linear controller, we add extra safety constraints to the LQR optimization problem (8), which leads to the following optimization problem.

$$\max_{M,Y} \quad Tr(W^{-1}Y) \tag{10a}$$

$$\text{s.t.} \quad \begin{bmatrix} -Y & Y & M^T & (AY+BM)^T \\ * & -Q^{-1} & 0 & 0 \\ * & * & -R^{-1} & 0 \\ * & * & * & -\frac{Y}{\gamma} \end{bmatrix} \preceq 0, \tag{10b}$$

$$Y \succeq 0, \tag{10c}$$

$$a_s^T Y a_s \leq 1, \quad s = 1, \cdots, q, \tag{10d}$$

This approach is, however, restrictive because using a linear controller to achieve both safety and optimality can quickly lead to infeasibility. This is because there typically exists no controller that can achieve optimal performance in an entire safety set while respecting safety. The conflict between safety and optimality results in infeasibility. To resolve the conflict, one can relax the optimality criteria to some extent using the LQR dual formulation (9). However, one cannot add safety constraint to the LQR dual problem as W_c is not related to the safety directly.

3.2 Indirect Safe LQR

In this approach, a controller is found that solves the first phase of Problem 2. While there is no general solution to it, one can find a controller that reaches a conflict-free zone for a large subset of the safe set instead of the entire safe set. This will resolve the infeasibility issue of the optimization problem . To this end, we parameterize the control gain K using safety specifications in the form of $\lambda-$contractivity. In other words, for a given polyhedral safe set, first, we parameterize the control gain K by a conflict-free zone ellipsoidal $\lambda-$contractive set P_s inside the original safe polyhedral set, then substitute that gain into the LQR dual problem to optimize the cost function. The conflict-free zone is formed by imposing safety and optimality conditions on a subset of the safe set and its size is maximized. Finally, we redefine the optimization objective to balance the trade-off between the conflict level and optimality by maximizing the size of the conflict-free set P_s besides optimizing the LQR cost function. We may lose a degree of performance to achieve some reasonable safety.

Theorem 1. Consider the system (1) with known A and B that is controlled by (2). Let Assumptions 1 and 3 hold and x_0 be a given initial condition. Then, the safety-parameterized optimal state feedback matrix K solves the safe LQR problem with known system matrices (*Phase 1 of Problem 2*) is given by

$$K = FP_s^{-1}, \tag{11}$$

where $P_s \succ 0$ and $F \in \mathbb{R}^{m\times n}$ are the solutions (if they exist) to the following SDP optimization problem

$$\min_{\beta>0,F,P_s\succeq 0,W_c\succeq W,H,L} \quad \alpha_1\beta - \alpha_2 \log\det P_s \tag{12a}$$

$$\text{s.t.} \quad \begin{bmatrix} W_c - W & AP_s + BF \\ * & \frac{H}{\gamma} \end{bmatrix} \succeq 0, \tag{12b}$$
$$\begin{bmatrix} \lambda P_s & (AP_s + BF)^T \\ * & P_s \end{bmatrix} \succeq 0, \tag{12c}$$
$$\begin{bmatrix} H & P_s \\ * & W_c \end{bmatrix} \succeq 0, \begin{bmatrix} L & F \\ * & H \end{bmatrix} \succeq 0, \tag{12d}$$
$$a_s^T P_s a_s \leq 1, \quad s = 1, \cdots, q, \tag{12e}$$
$$tr(QW_c) + tr(RL) \leq \beta, \tag{12f}$$

where $\alpha_1, \alpha_2 > 0$ are design parameters to tune the safety-optimality balance. Moreover, the set $\{x : x^T P_s x \leq 1\}$ provides a conflict-free zone.

Proof. Let us start by parameterizing the policy gain K considering the $\lambda-$contractivity of the ellipsoidal safe set (in expectation) characterized by $V(x_k) = x_k^T P_s^{-1} x_k$ inside the given polyhedral safe set $\mathcal{P}$ (4) as

$$\mathbb{E}[V(x_{k+1})|x_k] \leq \lambda V(x_k)$$
$$(A + BK)^T P_s^{-1}(A + BK) - \lambda P_s^{-1} \preceq 0$$
$$(AP_S + BF)^T P_s^{-1}(AP_s + BF) - \lambda P_s \preceq 0, \tag{13}$$

where $\lambda \in (0, 1)$ is the contractivity factor of the safe set defined by P_s and $F = KP_s$. The inequality (13) can be written as (12c) using the Schur-complement lemma (Boyd et al., 1994). We can find the maximum $\lambda-$contractive set identified by P_s inside the given safe set $\mathcal{P}$ by solving the following optimization problem

$$\min_{F,P_s} \quad -\log\det P_s \tag{14}$$
$$\text{s.t.} \quad \begin{bmatrix} \lambda P_s & (AP_s + BF)^T \\ * & P_s \end{bmatrix} \succeq 0,$$
$$a_s^T P_s a_s \leq 1, \quad s = 1, \cdots, q.$$

Now, consider the LQR dual problem (9) and substitute the safety-parameterized policy gain (11) in (9b) and (9c) as

$$\gamma(AP_s + BF)P_s^{-1}W_c P_s^{-1}(AP_s + BF)^T - W_c + W \preceq 0, \tag{15}$$
$$L - FP_s^{-1}W_c P_s^{-1}F^T \succeq 0. \tag{16}$$

Define a new matrix H such that

$$P_s^{-1}W_c P_s^{-1} \succeq H^{-1}. \tag{17}$$

This matrix shows the connection between the optimal state covariance matrix (controllability Gramian) W_c and the $\lambda-$contractive safe set defined by P_s^{-1}. The optimality criteria (9) tends to minimize the weighted trace of W_c and hence the trace of W_c. On the other hand, the safety criteria (14) increases the determinant of P_s and hence the size of P_s. So, the H matrix shows the conflict between safety and optimality. Also, it reveals that safety is more dominant than optimality in this method. As a result, we introduced two positive weight parameters α_1 and α_2 to let the designer balance optimality and safety. Since the safety is dominant, α_2 should be chosen as small as possible.

4. DATE-DRIVEN SAFE LQR

This section extends the previous results to the data-driven scenarios. Safety is of paramount importance especially when we are designing a direct data-driven (model-free) LQR.

Assumption 4. The collected data matrix $D_0 = \begin{bmatrix} U_0^T & X_0^T \end{bmatrix}^T$ has full row rank, i.e. $rank(D_0) = m + n$.

Consider the following parameterization of the closed-loop system (De Persis and Tesi, 2021) described by the collected data, safety matrix, and a matrix $M \in \mathbb{R}^{N\times n}$

$$A + BK = (X_1 - \Omega_0)MP_s^{-1}, \tag{18a}$$
$$X_0 M = P_s, \tag{18b}$$
$$K = U_0 M P_s^{-1}. \tag{18c}$$

The above equations show how the safety set depends on the collected data and its quality. The following theorem, which is the data-driven counterpart of Theorem 1, highlights the necessary conditions to design a safe optimal control that solves the first phase of *Problem 2.*

Theorem 2. Consider the system (1) with closed-loop data-based safety-aware parameterization (18). Let Assumptions 1-4 hold and x_0 be a given initial condition. Then, the safety-parameterized optimal state feedback matrix K that solves the safe LQR problem (*phase 1 of Problem 2*) can be computed through the following SDP optimization problem

$$\min_{\beta>0,\delta>0,M,P_s\succeq 0,W_c\succeq W,H,L} \quad \alpha_1\beta - \alpha_2 \log\det P_s \tag{19a}$$
$$\text{s.t.} \quad \begin{bmatrix} W_c - W & X_1 M \\ * & \frac{H}{\gamma} \end{bmatrix} \succeq 0, \tag{19b}$$
$$\begin{bmatrix} \lambda P_s & (X_1 M)^T & M^T \\ * & * & 0 \\ * & * & \delta \end{bmatrix} \succeq 0, \tag{19c}$$
$$tr(P_s W^{-1}) - \delta n^2 > 0, \tag{19d}$$
$$\begin{bmatrix} H & P_s \\ * & W_c \end{bmatrix} \succeq 0, \begin{bmatrix} L & F \\ * & H \end{bmatrix} \succeq 0, \tag{19e}$$
$$a_s^T P_s a_s \leq 1, \quad s = 1, \cdots, q, \tag{19f}$$
$$X_0 M = P_s, \tag{19g}$$
$$tr(QW_c) + tr(RL) \leq \beta, \tag{19h}$$

where $\alpha_1, \alpha_2 > 0$ are design parameters to tune the safety-optimality balance.

Proof. Let us consider the data-driven $\lambda-$contractivity in expectation of the ellipsoidal safe set defined by P_s inside the given polyhedral safety constraint (4) using the representation outlined in (18) as

$$\mathbb{E}[V(x_{k+1}|x_k] \leq \lambda V(x_k)$$
$$\mathbb{E}[P_s^{-1}M^T(X_1 - \Omega_0)^T P_s^{-1}(X_1 - \Omega_0)MP_s^{-1}] - \lambda P_s^{-1} \preceq 0$$
$$P_s^{-1}M^T X_1^T P_s^{-1} X_1 M P_s^{-1} + \mathbb{E}[P_s^{-1}M^T \Omega_0^T P_s^{-1}\Omega_0 M P_s^{-1}]$$
$$- \lambda P_s^{-1} \preceq 0$$
$$P_s^{-1}M^T X_1^T P_s^{-1} X_1 M P_s^{-1} + P_s^{-1}M^T tr(P_s^{-1}W)MP_s^{-1}$$
$$- \lambda P_s^{-1} \preceq 0. \tag{20}$$

By multiplying the inequality (20) from left and right by P_s, using the Schur-complement twice, and the following inequalities

$$\frac{1}{tr(W^{-1}P_s)} < \frac{tr(P_s^{-1}W)}{n^2}, \tag{21}$$
$$tr(P_s^{-1}W) \leq \frac{1}{\delta} \tag{22}$$

the data-driven $\lambda-$contractivity LMI (19c) and inequality (19d) will be resulted, where $\delta > 0$ is a decision variable.

The term $M^T tr(P_s^{-1}W)M$ in (20) brings robustness into the design of the data-driven controller. Now, consider the LQR dual problem (9) and substitute the approximate data-driven safety-parameterized closed-loop

$$A + BK \approx X_1 M P_s^{-1}, \tag{23}$$

in (9b) which results in

$$\gamma X_1 M H^{-1} M^T X_1^T - W_c + W \preceq 0. \tag{24}$$

This inequality is called certainty-equivalent data-driven condition (De Persis and Tesi, 2021) since we have no access to the noise matrix Ω_0 and ignore it. The data-driven LQR LMI (24) suffers the robustness feature since we disregarded the noise data, but the data-driven $\lambda-$contractivity LMI in expectation (20) compensates for the conscious ignorance of robustness in the optimality LMI. The rest of the proof is similar to the first theorem.

In the data-driven situation, since we have no accurate information about the system, we can choose a bigger value for α_2 compared to the model-based design to encourage safety while preserving the notion of optimality. Also, the maximum safe zone P_s depends on the quality of the collected data and λ.

5. SIMULATION RESULTS

This section compares the presented $\lambda-$Contractive LQR ($\lambda-$CLQR) (both model-based (12) and data-driven (19) derivatives) with the model-based LQR (8), the direct data-driven approach (Low Complexity LQR (LCLQR) and robust LCLQR) proposed in (De Persis and Tesi, 2021), and the $\lambda-$Contractive Control ($\lambda-$CC) (14). Specifically, we will compare their costs and safety violations. We used CVXPY (Diamond and Boyd, 2016; Agrawal et al., 2018) for modeling the convex optimization problems and MOSEK (ApS, 2022) as the solver.
Consider the vehicle steering model (Kishida and Cetinkaya, 2022)

$$A = \begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0.06 \\ 0.2 \end{bmatrix}, W = \begin{bmatrix} 0.01 & 0.003 \\ 0.003 & 0.02 \end{bmatrix}$$

which describes the lateral deviation dynamics with x_1 and x_2 being the lateral position and heading angle, respectively. The LQR matrices are chosen as $Q = 10I_2$ and $R = 1$, and the contractivity factor is $\lambda = 0.91$. Fig. 1 shows the trajectories obtained from the model-based LQR and $\lambda-$CC designs. On the one hand, the LQR violates the safety presented as the dash-dotted green rectangle, while $\lambda-$CC preserves the safety. On the other hand, the cost of LQR is much less than the $\lambda-$CC. Fig. 2 represents the trajectories obtained by the $\lambda-$CLQR. Safety violations using the presented approach are significantly decreased compared to the LQR. This is because the presented approach accounts for conflict awareness in the design time, which can be leveraged to significantly decrease the CBC interventions in the real-time. This is shown in the next two figures. Figs. 3 and 4 show the trajectories of the data-driven $\lambda-$CLQR and LCLQR for two sets of collected data. In Fig. 3, there is no CBC mechanism, so LCLQR has no safety awareness and its trajectories violate the safety region. At the same time, $\lambda-$CLQR preserves safety with a small increase in cost. Another feature of $\lambda-$CLQR is that it provides an estimate of the safe region for the system's operation based on the collected data, as it is obvious

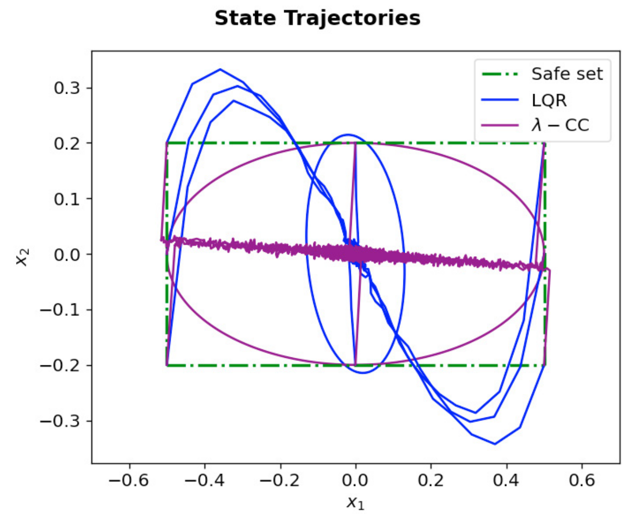

Fig. 1. Model-based controllers, LQR: cost= 0.0115, λ-CC: cost=0.1200.

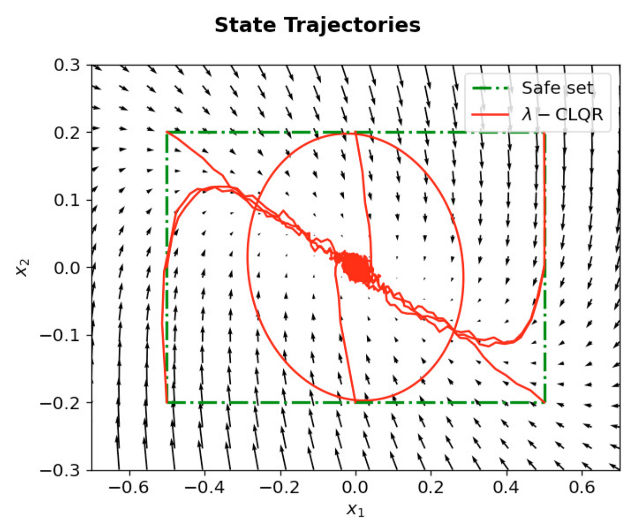

Fig. 2. Model-based controller, λ-CLQR: cost=0.0217, $\alpha_1 = 4000, \alpha_2 = 0.00005$.

from the red ellipsoidal in Figs. 3 and 4. In Fig. 4, the CBC component is used with system matrices estimated from the collected data (which are not exact). Therefore, even with the CBC mechanism, LCLQR can not provide safety, while $\lambda-$CLQR, which has the safety-awareness feature, can maintain the system's safety with a slight increase in cost due to the CBC interventions. This clearly shows the advantage of the presented conflict-aware safe LQR. Fig. 5 shows the performance and safety awareness of the robust version of LCLQR with $\alpha = 0.01$ and $\lambda-$CLQR with the CBC interventions. The performance of $\lambda-$CLQR is better than the LCLQR as it requires 51 CBC interventions compared to 152 CBC interventions of the LCLQR. It shows how the controller's safety awareness can help reduce costs while maintaining safety.

6. CONCLUSION

In conclusion, this paper has presented a novel approach to integrating safety considerations into the design of LQR

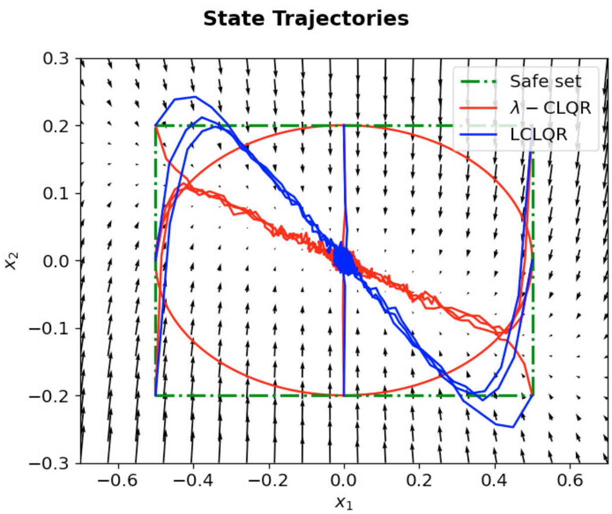

Fig. 3. Data-driven controllers, λ-CLQR: $\alpha_1 = 4000, \alpha_2 = 5$, cost=0.024, LCLQR: cost=0.0132

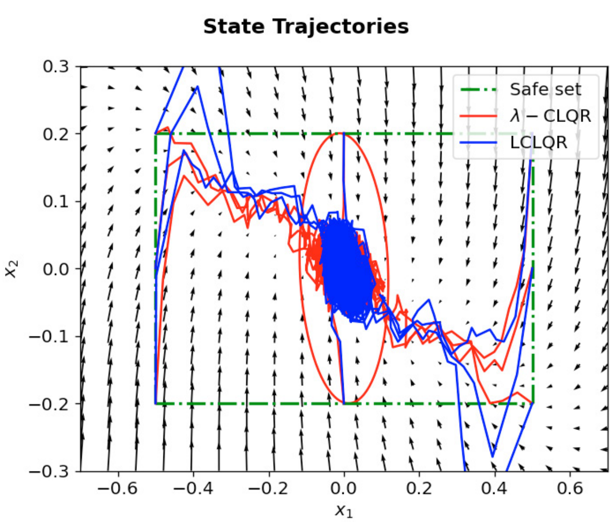

Fig. 4. Data-driven controllers with CBC, quality of data and its effect on the maximum achievable safe set, λ-CLQR: $\alpha_1 = 4000, \alpha_2 = 5$, cost=0.029, LCLQR: cost=0.026

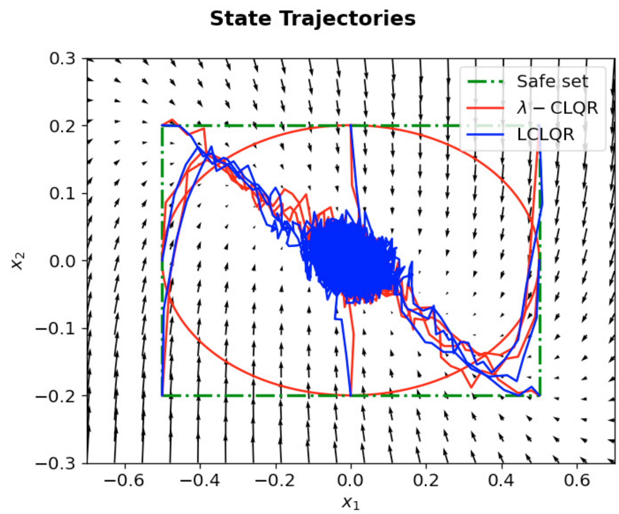

Fig. 5. Data-driven controllers with CBC, $\alpha_1 = 4000, \alpha_2 = 5$, λ-CLQR cost=0.029, Robust LCLQR: cost=0.038, $\alpha = 0.01$.

controllers within the framework of direct data-driven control by leveraging the λ-contractivity principle. Extending the proposed framework to incorporate more sophisticated safety constraints and nonlinear system dynamics would be an interesting direction for future research.

REFERENCES

Agrawal, A., Verschueren, R., Diamond, S., and Boyd, S. (2018). A rewriting system for convex optimization problems. *Journal of Control and Decision*, 5(1), 42–60.

Agrawal, A. and Sreenath, K. (2017). Discrete control barrier functions for safety-critical control of discrete systems with application to bipedal robot navigation. In *Robotics: Science and Systems*, volume 13, 1–10. Cambridge, MA, USA.

Ames, A.D., Coogan, S., Egerstedt, M., Notomista, G., Sreenath, K., and Tabuada, P. (2019). Control barrier functions: Theory and applications. In *2019 18th European control conference (ECC)*, 3420–3431. IEEE.

ApS, M. (2022). *MOSEK Optimizer API for Python 9.3.22*. URL `https://docs.mosek.com/9.3/pythonapi/index.html`.

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.

Choi, J., Castaneda, F., Tomlin, C.J., and Sreenath, K. (2020). Reinforcement learning for safety-critical control under model uncertainty, using control Lyapunov functions and control barrier functions. *arXiv preprint arXiv:2004.07584*.

De Persis, C. and Tesi, P. (2021). Low-complexity learning of linear quadratic regulators from noisy data. *Automatica*, 128, 109548.

Diamond, S. and Boyd, S. (2016). CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83), 1–5.

Hou, Z.S. and Wang, Z. (2013). From model-based control to data-driven control: Survey, classification and perspective. *Information Sciences*, 235, 3–35.

Kishida, M. and Cetinkaya, A. (2022). Risk-aware linear quadratic control using conditional value-at-risk. *IEEE Transactions on Automatic Control*, 68(1), 416–423.

Krishnan, V. and Pasqualetti, F. (2021). On direct vs indirect data-driven predictive control. In *2021 60th IEEE Conference on Decision and Control (CDC)*, 736–741. doi:10.1109/CDC45484.2021.9683187.

Lee, D. (2023). Multi-objective lqg design with primal-dual method. *Mathematics*, 11(8).

Markovsky, I. and Dörfler, F. (2021). Behavioral systems theory in data-driven analysis, signal processing, and control. *Annual Reviews in Control*, 52, 42–64.

Zanon, M. and Gros, S. (2020). Safe reinforcement learning using robust MPC. *IEEE Transactions on Automatic Control*, 1–1.