Conformal Robust Control of Linear Systems

Yash Patel
University of Michigan

Sahana Rayan
University of Michigan

Ambuj Tewari
University of Michigan

Abstract

End-to-end engineering design pipelines, in which designs are evaluated using concurrently defined optimal controllers, are becoming increasingly common in practice. To discover designs that perform well even under the misspecification of system dynamics, such end-to-end pipelines have now begun evaluating designs with a robust control objective. Current approaches of specifying such robust control subproblems, however, rely on hand specification of perturbations anticipated to be present upon deployment or margin methods that ignore problem structure, resulting in a lack of theoretical guarantees and overly conservative empirical performance. We, instead, propose a novel methodology for LQR systems that leverages conformal prediction to specify such uncertainty regions in a data-driven fashion. Such regions have distribution-free coverage guarantees on the true system dynamics, in turn allowing for a probabilistic characterization of the regret of the resulting robust controller. We demonstrate that such a controller can be efficiently synthesized via a novel policy gradient method that has convergence guarantees and that this robust controller empirically outperforms alternate robust control methods, such as $\mathcal{H}_\infty$ and LQR with multiplicative noise, across several engineering tasks.

1 INTRODUCTION

Seeking control over a family of dynamical systems is a problem often encountered in engineering (Killian, Konidaris, and Doshi-Velez, 2016; Wu et al., 2018; Aksland et al., 2023). One prevalent application of this is in cases where engineering designs and their respective controllers are being concurrently developed, known as control co-design (CCD) (Garcia-Sanz, 2019). Traditional engineering design loops operated sequentially, first proposing a design and then developing a controller (Reyer et al., 2001; Friedland, 1995). Such workflows, however, sacrificed the improved optimality possible in their coupling, hence the increasing interest in leveraging end-to-end co-control design pipelines (Fathy et al., 2001; Falck et al., 2021).

Initial works in CCD studied optimal design assuming perfectly specified, deterministic system dynamics (Allison, Guo, and Han, 2014; Azad et al., 2018; Azad et al., 2019; Behtash and Alexander-Ramos, 2018). Such assumptions have, however, become overly restrictive, resulting in interest in robust extensions of the CCD formulation, referred to as uncertain CCD (UCCD) (Azad and Herber, 2022; Azad and Herber, 2023a; Bird et al., 2023). Such uncertainty can arise from many sources in the design process, such as noise in the controllers, uncertainties in the design parameters, or unmodeled dynamics. The UCCD specification also differs depending on the risk tolerance in the downstream application. For instance, in risk-neutral settings, stochastic specifications are appropriate (Azad and Alexander-Ramos, 2020a; Cui, Zheng, and Wang, 2022; Behtash and Alexander-Ramos, 2024), whereas in risk-averse settings, probabilistic (Cui, Allison, and Wang, 2020a; Cui, Allison, and Wang, 2020b; Nguyen, Kuo, and Lin, 2022) or worst-case forms (Azad and Alexander-Ramos, 2021; Nash, Pangborn, and Jain, 2021; Azad and Alexander-Ramos, 2020b) are used.

We focus on the worst-case robust UCCD formulation (WCR-UCCD), specifically on dynamics misspecification. WCR-UCCD requires specifying a dynamics uncertainty region. Existing methods of specification, however, tend to be ad-hoc and, thus, fail to provide any guarantees of the robust solution as it relates to the selection of this uncertainty set, rendering its choice often difficult and resulting in suboptimal controller synthesis (Azad and Alexander-Ramos, 2020a).

We, thus, focus herein on providing a principled

Proceedings of the 29[th] International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300.

distribution-free specification of the *robust control subproblem* in WCR-UCCD and an associated solution method with convergence guarantees. One special case of interest in UCCD is in the setting of linear quadratic regulators (LQRs), where the underlying system dynamics take on a linear structure (Ahmadi, Rahmani, and Shahmansoorian, 2023; Fathy et al., 2003; Jiang et al., 2016). LQR systems are of broad interest both due to their analytic tractability and widespread applicability to practical engineering systems (Zhao et al., 2024; Mamakoukas et al., 2019; Bevanda et al., 2022). We, therefore, propose a method for specifying the LQR WCR-UCCD control subproblem that lends itself to efficient solution by leveraging conformal prediction on observed design information. A related use of conformal prediction for predict-then-optimize problems was recently studied in (Patel, Rayan, and Tewari, 2024). Unlike their setting, however, the application of conformal prediction to control has complications related to the stability of controllers under model uncertainty. Our contributions are as follows:

- Providing a framework to define robust LQR control problems with distribution-free probabilistic regret guarantees, across deterministic or stochastic and discrete- or continuous-time dynamics, and demonstrating empirical improvements over alternative robust control schemes.
- Extending conformalized predict-then-optimize to cases where calibration data is observed with noise and where the domains of both the maximization *and* minimization components of the robust formulation depend on the conformalized predictor.
- Providing a novel policy subgradient method for robust controller synthesis with convergence guarantees proven via subgradient dominance.

2 BACKGROUND

2.1 Conformal Prediction

Conformal prediction is a principled, distribution-free approach of uncertainty quantification (Angelopoulos and Bates, 2021; Shafer and Vovk, 2008). "Split conformal," the most common variant of conformal prediction, is used as a wrapper around predictors $\widehat{f} : \mathcal{X} \to \mathcal{Y}$ such that prediction *regions* $\mathcal{U}(x)$ are returned in place of the point predictions $\widehat{f}(x)$. Prediction regions $\mathcal{U}(x)$ are sought to have coverage guarantees on the true $y := f(x)$. That is, for some prespecified α, we wish $\mathcal{P}_{X,Y,\mathcal{U}}(Y \in \mathcal{U}(X)) \geq 1 - \alpha$, where randomness in $\mathcal{U}$ may arise from algorithmic choices.

To achieve this, split conformal partitions the overall dataset $\mathcal{D}$ into two subsets, $\mathcal{D}_T \cup \mathcal{D}_C$, respectively the training and calibration datasets. After fitting $\widehat{f}$ on the training subset, the calibration set is used to measure the anticipated "prediction error" for future test points. Formally, this error is quantified via a score function $s(x, y)$, which generalizes the classical notion of a residual. In particular, scores are evaluated on the calibration dataset to define $\mathcal{S} := \{s(x, y) \mid (x, y) \in \mathcal{D}_C\}$. Denoting the $\lceil (|\mathcal{D}_C| + 1)(1 - \alpha) \rceil / |\mathcal{D}_C|$ empirical quantile of $\mathcal{S}$ as $\widehat{q}$, conformal prediction defines $\mathcal{U}(x) := \{y \mid s(x, y) \leq \widehat{q}\}$. Such $\mathcal{U}(x)$ satisfies $\mathcal{P}_{X,Y,\widehat{q}}(Y \in \mathcal{U}(X)) \geq 1 - \alpha$ under the exchangeability of future test points (x', y') with points from $\mathcal{D}_C$. While the coverage guarantee holds for any $s(x, y)$, the sizes of the resulting prediction regions, known as the procedure's "predictive efficiency," is dependent on its choice (Shafer and Vovk, 2008).

2.2 Robust Predict-Then-Optimize

Predict-then-optimize problems are nominally formulated as $w^*(x) := \min_{w \in \mathcal{W}} \mathbb{E}[f(w, C) \mid x]$, where w are decision variables, C an *unknown* cost parameter, x observed contextual variables, $\mathcal{W}$ a compact feasible region, and $f(w, c)$ an objective function that is convex-concave and L-Lipschitz in c for any fixed w. The nominal approach defines a predictor $\widehat{g} : \mathcal{X} \to \mathcal{C}$, where the prediction $\widehat{c} := \widehat{g}(x)$ is leveraged, i.e. taking $w^*(x) := \min_w f(w, \widehat{c})$. Such an approach, however, is inappropriate in safety-critical settings, given that $\widehat{g}$ will likely be misspecified and, thus, result in decisions that are suboptimal under the true cost parameter, c. For this reason, robust alternatives to the nominal formulation have become of interest (Chenreddy, Bandi, and Delage, 2022; Sadana et al., 2024; Chenreddy and Delage, 2024). We focus on the following formulation from (Patel, Rayan, and Tewari, 2024):

$$w^*(x) := \min_{w} \max_{\widehat{c} \in \mathcal{U}(x)} \quad f(w, \widehat{c}) \tag{1}$$

where $\mathcal{U} : \mathcal{X} \to \mathcal{F}$ is a uncertainty region predictor, with $\mathcal{F}$ being the σ-field of $\mathcal{C}$, such that $\mathcal{P}_{X,C,\mathcal{U}}(C \in \mathcal{U}(X)) \geq 1 - \alpha$. Works in this field typically study the suboptimality gap, defined as $\Delta(x, c) := \min_w \max_{\widehat{c} \in \mathcal{U}(x)} f(w, \widehat{c}) - \min_w f(w, c)$. For instance, in (Patel, Rayan, and Tewari, 2024; Johnstone and Cox, 2021; Sun, Liu, and Li, 2023; Yeh et al., 2024), $\mathcal{U}(x)$ was constructed via conformal prediction to provide probabilistic guarantees; that is, by taking $\mathcal{U}(x)$ to be the prediction region produced by conformalizing the predictor $\widehat{g}$, they demonstrated $\mathcal{P}_{X,C,\widehat{q}}(0 \leq \Delta(X, C) \leq L \operatorname{diam}(\mathcal{U}(X))) \geq 1 - \alpha$.

2.3 LQR & Control Co-Design

The field of control has a long history in engineering physics and robotics (Zabczyk, 2020). In the linear

quadratic regulator (LQR) setup, the state dynamics have a linear form. Notably, system dynamics can be studied in both continuous- and discrete-time. For clarity, we focus the exposition of this manuscript on the *discrete*-time setting, though many of our presented results extend immediately to the continuous-time setting. Where appropriate, we have included such continuous-time results in the Appendix.

A discrete-time linear system evolves as $x_{t+1} = Ax_t + Bu_t + w_t$, where x_t is the state at time t, u_t the control input at time t, and $w_t \sim \mathcal{D}_w$ the noise. Optimal control is then posed as an optimization problem, with the objective $J(u)$ weighing both the deviation from a target state and the necessary control input. LQR optimal controllers take a linear feedback form, namely $u_t^* = -K^* x_t$ where K^* is known as the "optimal gain matrix" and solves

$$K^*(A,B) := \underset{K \in \mathcal{K}(A,B)}{\arg\min}\ \mathbb{E}[J(K,A,B)] \qquad (2)$$
$$\text{where } J(K,A,B) := \sum_{t=0}^{\infty} (x_t^\top Q_t x_t + (Kx_t)^\top R_t (Kx_t))$$

where $Q_t = Q_t^\top \succcurlyeq 0$ and $R_t = R_t^\top \succcurlyeq 0$. Q_t and R_t model the user's preference of state deviation and control input and can vary with t, though non-varying $Q_t = Q$ and $R_t = R$ are often taken. $\mathcal{K}(A,B) := \{K : \rho(A - BK) < 1\}$ is the set of "stabilizing controllers" for the A, B system dynamics, where $\rho(\cdot)$ denotes the spectral radius. Solving this is done either with the algebraic Riccati equation (ARE) (Willems, 1971) or via policy gradient (Sun and Fazel, 2021).

We now briefly discuss uncertain co-control design; for a full survey, refer to (Azad and Herber, 2022). Engineering designs can often be specified by parameters θ, which could capture, for instance, the dimensions of an airfoil or material properties of a DC battery grid. The dynamics are highly dependent on the design; for example, an airfoil with a shape θ_1 will fly differently from one given by θ_2. Worst-case robust UCCD with dynamics misspecification, thus, solves $\min_{K,\theta} \max_{\widehat{A} \in \mathcal{A}(\theta), \widehat{B} \in \mathcal{B}(\theta)} \mathbb{E}[o(K, \widehat{A}, \widehat{B}, \theta)]$, where $x_{t+1} = \widehat{A}x_t + \widehat{B}u_t + w_t$ and $(\mathcal{A}(\theta), \mathcal{B}(\theta))$ are uncertainty sets of the dynamics for such a design and o is the objective.

Often, the objective takes a decomposable form, namely with one term relating to system control and the other depending on the design parameter, i.e. $o(K, A, B, \theta) := \ell(\theta) + J(K, A(\theta), B(\theta))$ (Chanekar, Chopra, and Azarm, 2018; Ahmadi, Rahmani, and Shahmansoorian, 2023). One commonly applied solution technique in this setting is bilevel optimization, in which an outer optimization loop is performed over design parameters and an inner one over controllers for the current design iterate (Herber and Allison, 2019; Kamadan, Kiziltas, and Patoglu, 2017). For this reason, the specification of the robust control subproblem can be studied independently of the outer design optimization loop, as done herein.

3 METHODOLOGY

We now discuss conformally robust LQR, providing the formulation in Section 3.1, regret guarantees in Section 3.3 and Section 3.4, and a controller synthesis algorithm with convergence guarantees in Section 3.5.

3.1 Problem Formulation

For the presentation below, let $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, $A \in \mathbb{R}^{n \times n}$, and $B \in \mathbb{R}^{n \times m}$. Let C denote the full dynamics matrix $C := [A, B] \in \mathbb{R}^{n \times (n+m)}$. We additionally assume a linear control scheme, namely $u_t = -Kx_t$ for some gain matrix K. Additionally, denote $W := [I_{n \times n} \ -K^\top]^\top \in \mathbb{R}^{(n+m) \times n}$, such that the closed-loop dynamics are given by $CW = A - BK$. As discussed in Section 2.3, we assume a dataset of designs and associated trajectories is observed. We assume such a dataset $\mathcal{D}$ consists of N samples $(\theta^{(i)}, C^{(i)}) \overset{\text{iid}}{\sim} \mathcal{P}(\Theta, C)$ and $K^{(i)} \overset{\text{iid}}{\sim} \mathcal{P}(K)$, where $\mathcal{P}(\Theta, C)$ is an unknown joint distribution over designs and dynamics and $\mathcal{P}(K)$ an unknown distribution on gain matrices. We make no assumptions on such distributions other than that each gain matrix $K^{(i)}$ is a stabilizing controller for the respective $C^{(i)}$ dynamics. Note that these underlying true dynamics $C^{(i)}$ are never observed directly by the learning algorithm; only the resulting trajectories are observed. Such trajectories are generated by evolving the state via $x_{t+1}^{(i)} = (C^{(i)} W^{(i)}) x_t^{(i)}$ over a time horizon T. The final dataset, therefore, takes the form $\mathcal{D} = \{\theta^{(i)}, \{(x_t^{(i)}, u_t^{(i)})\}_{t=1}^T\}_{i=1}^N$. We are interested in studying a risk-sensitive formulation of LQR:

$$K^*_{\text{rob}}(\mathcal{U}(\theta)) := \underset{K \in \mathcal{K}(\mathcal{U}(\theta))}{\arg\min} \max_{[\widehat{A}, \widehat{B}] := \widehat{C} \in \mathcal{U}(\theta)} \mathbb{E}[J(K, \widehat{A}, \widehat{B})]$$
$$\text{s.t.} \quad x_{t+1} = \widehat{A}x_t + \widehat{B}u_t + w_t$$
$$\mathcal{P}_{\Theta, C, \mathcal{U}}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha,$$

where J is the objective function particular to the setting of interest, differing between infinite and finite time horizons and continuous and discrete time dynamics, and $\mathcal{U}(\theta)$ is an uncertainty set over dynamics. Notably, the notion of stabilizing controllers must be generalized in this robust formulation, since the nominal formulation is for a specific C. We, thus, consider those controllers that stabilize the entire uncertainty set, which we refer to as the "universal stabilizing set," formally $\mathcal{K}(\mathcal{U}(\theta)) := \bigcap_{\widehat{C} \in \mathcal{U}(\theta)} \mathcal{K}(\widehat{C})$, where $\mathcal{K}(\widehat{C})$ is Equation (2) evaluated for a particular $\widehat{A}, \widehat{B}$.

3.2 Score Function

From the trajectories in $\mathcal{D}$, we can perform system identification using least squares estimation to recover estimates of the system dynamics, $(\widetilde{A}^{(i)}, \widetilde{B}^{(i)})$ (Ljung et al., 1987). With this, we obtain a final dynamics dataset $\widetilde{\mathcal{D}} = \{\theta^{(i)}, \widetilde{C}^{(i)}\}_{i=1}^{N}$, which we then leverage in the standard manner of split conformal prediction. That is, we split $\widetilde{\mathcal{D}} = \widetilde{\mathcal{D}}_{\mathcal{T}} \cup \widetilde{\mathcal{D}}_{\mathcal{C}}$, the former of which we use to train a system parameters predictor $\widehat{C} := f(\theta)$. Notably, leveraging split conformal in this setting has the complication that the ground truth used, namely in $\widetilde{\mathcal{D}}_{\mathcal{C}}$, is itself an estimate $\widetilde{C}$ even though coverage is sought on C. We assume for this initial discussion that for a fixed coverage level α, we can obtain prediction regions with the desired coverage, satisfying $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{U}(\Theta)) \geq 1-\alpha$, using $\widetilde{\mathcal{D}}_{\mathcal{C}}$. The treatment of this gap between $\widetilde{C}$ and C is discussed in Section 3.4.

We take the score to be $s(\theta, C) = ||f(\theta) - C||_{\text{op}}$, where $||\cdot||_{\text{op}}$ is the matrix *operator* norm, i.e. $||A||_{\text{op}} = \sigma_{\max}(A)$, from which the resulting prediction regions take on the form of $\mathcal{B}_{\widehat{q}}(f(\theta))$, namely a ball of radius $\widehat{q}$, the conformal quantile, under the $||\cdot||_{\text{op}}$ metric.

3.3 Coverage Guarantee Consequences

We now characterize the regret induced by the robustness across LQR setups, that is $\mathcal{R}(\theta, C) := \mathbb{E}[J(K^*_{\text{rob}}(\mathcal{U}(\theta)), C) - J(K^*(C), C)]$, where the randomness is over stochastics in the *true* system dynamics $C := [A, B]$ and in the $\mathcal{P}(C \mid \theta)$ map. We explicitly note C in the regret notation to emphasize that, while the controller K is defined using *estimated* system dynamics, the final evaluation is over the *true* C.

As mentioned previously, while we present the results for the discrete-time setting below for clarity, the continuous-time results follow analogously and are presented fully in the Appendix. We present below both the cases of deterministic and stochastic dynamics. Both settings require a mild assumption that the problem parameters have bounded norms, formalized in Assumption 3.1; this will hold for any realistic problem setup. Notably, however, the two settings differ in that the stochastic dynamics requires Q_t and R_t be discounted over t, while the deterministic case is fully compatible with non-discounted rewards. Intuitively, this discounting is necessary, as stability alone in the stochastic setting does not ensure a bounded objective; the state can continue to oscillate and result in an unbounded accumulation of error if the terms tied to the state covariance matrix do not decay. Other works frame this assumption as "mean-square stability," (see e.g. Gravell, Esfahani, and Summers, 2020a). We formally pose this as Assumption 3.2.

Assumption 3.1. $\forall\ \theta,\ K \in \mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$,

$$D(K) := \max_{t \geq 0} \sqrt{n} ||Q_t + K^\top R_t K||_\infty ||x_0||_\infty^2 ||W||_{\text{op}} < \infty$$

Assumption 3.2. For any θ, $\exists$ constants $\alpha_1, \beta_1 > 0$ such that for all $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$, $K \in \mathcal{K}(\widehat{C})$, and $t \geq 0$, $||Q_t + K^\top R_t K|| \leq \beta_1 e^{-\alpha_1 t}$ and $\min_{\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))}(2\alpha_2(\widehat{C}) + \alpha_1) > 0$ where $\alpha_2(\widehat{C}) := \inf_{K \in \mathcal{K}(\widehat{C})}(-\log \rho(\widehat{C}W)) > 0$.

The regret bound below decomposes into two terms. The first captures the suboptimality in designing a controller with the conformal dynamics set instead of against the true dynamics; this coincides with the suboptimality characterized in previous works described in Section 2.2. The other is a novel aspect that arises in this controls setting: since the robust control problem optimizes over a *restricted* set of controllers, namely those that universally stabilize the full conformal dynamics set instead of those that only stabilize the true dynamics, there is an additional "domain gap" term.

Definition 3.3 (Domain Gap)**.** Let $\mathcal{B}_{\widehat{q}}(f(\theta))$ be as defined in Section 3.1 for a true dynamics C. Then,

$$\Delta_{\text{dom}}(\theta, C) := |\min_{K \in \mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} J(K, C) - \min_{K \in \mathcal{K}(C)} J(K, C)|$$

Intuitively, if the true optimal controller falls in $\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$, this latter term should vanish. Towards this end, we introduce the following notion.

Definition 3.4 (Radius Threshold)**.** Let $M(C, K^*(C)) := A - BK^*(C)$ be diagonalizable. Let

$$r(C, K^*(C)) := \frac{\min_i(1 - |\lambda_i(M)|)}{2\kappa(U)\,||W||_{\text{op}}},$$

where $M = U\Lambda U^{-1}$, $\kappa(U)$ is the condition number of U, and $W = [I\ \ -K^*(C)^\top]^\top$.

Across the theorems stated below, therefore, if the conformal radius is smaller than this $r(C, K^*(C))$ threshold, the "domain gap" term vanishes. Intuitively, this property follows as the stability of the system can be characterized by the closed-loop eigenvalues, whose values change by a bounded amount in considering the perturbations captured in the conformal region. We additionally see that, as $\widehat{q} \to 0$, the suboptimality vanishes, as we would expect in recovering the true dynamics. Thus, users should seek to produce prediction regions with coverage that are as small as possible to produce informative upper bounds on the nominal optimal value. Notably, the statements below are given in terms of the objective Lipschitz constants L: explicit expressions of L along with the continuous-time and finite time horizons theorems and proofs are given in Appendices B to E.

Theorem 3.5 (Deterministic)**.** *Let* $J(K,C) := \sum_{t=0}^{\infty}(x_t^\top(Q_t + K^\top R_t K)x_t)$ *with* $w = 0$*. Assume that* $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\Theta))) \geq 1-\alpha$*. Then, under Assumption 3.1,*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\left(0 \leq \mathcal{R}(\Theta,C) \leq 2L\widehat{q} + \Delta_{\text{dom}}(\Theta,C)\right) \geq 1-\alpha,$$

where L *is the Lipschitz constant of* $J(K,\widehat{C})$ *in* $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ *under the operator norm. Further, if* $\widehat{q} < r(C, K^*(C))$*,* $\Delta_{\text{dom}}(\Theta,C) = 0$*, as defined in Definitions 3.3 and 3.4.*

Theorem 3.6 (Stochastic)**.** *Let* $J(K,C) := \sum_{t=0}^{\infty}(x_t^\top(Q_t + K^\top R_t K)x_t)$ *with* $w_t \overset{\text{iid}}{\sim} \mathcal{N}(0,\Sigma)$ *across* t *such that* $D_2(K) := ||\Sigma||_{\text{op}}||W||_{\text{op}} < \infty$*. Assume further that* $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\Theta))) \geq 1-\alpha$*. Then, under Assumption 3.1 and Assumption 3.2,*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\left(0 \leq \mathcal{R}(\Theta,C) \leq 2L\widehat{q} + \Delta_{\text{dom}}(\Theta,C)\right) \geq 1-\alpha,$$

where L *is the Lipschitz constant of* $J(K,\widehat{C})$ *in* $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ *under the operator norm. Further, if* $\widehat{q} < r(C, K^*(C))$*,* $\Delta_{\text{dom}}(\Theta,C) = 0$*, as defined in Definitions 3.3 and 3.4.*

3.4 Ambiguous Ground Truth

We now discuss the complication of obtaining coverage guarantees on C despite only observing estimates $\widetilde{C} = C + \epsilon$ in the dataset, where $\epsilon \sim \mathcal{N}(0,\Sigma)$. This form of the estimation error can be shown to hold asymptotically under mild assumptions by classical results from least squares estimation, as shown for LTI system identification in (Ljung et al., 1987).

The coverage guarantee result given in Theorem 3.7 is the multivariate extension of Theorem A.5 from (Feldman et al., 2023) and is a novel contribution to the broader space of conformal prediction. Intuitively, we show that if, for all θ, the density $\mathcal{P}(C \mid \theta)$ peaks in $\mathcal{U}(\theta)$, we retain marginal coverage guarantees. If $\mathcal{P}(C \mid \theta)$ is unimodal and radially symmetric about its mode, this condition is satisfied so long as $\mathcal{U}(\theta)$ captures the mode. The map between design parameters θ and A, B is often unimodal, making such a structural assumption reasonable; this was true classically, where a deterministic map was parametrically given by physics (discussed more in Appendix F), and remains true of data-driven surrogates in UCCD (Azad and Herber, 2023b; Azad et al., 2024). $\mathcal{U}(\theta)$ capturing the mode is also a weak assumption assuming a zero-centered distribution for ϵ, since it then amounts to capturing the mode of $\mathcal{P}(\widetilde{C} \mid \theta)$, which holds for any sufficiently accurate predictor. We empirically demonstrate that such assumptions hold and, thus, that the coverage guarantees are retained in Section 5. The full proof of this theorem is deferred to Appendix G.

Theorem 3.7. *Let* $\widetilde{C} = C + \epsilon$ *where* $vec(\epsilon) \sim \mathcal{N}(0,\Sigma)$*, where* $\epsilon \perp\!\!\!\perp (\Theta, C)$*. Assume* $\mathcal{U}(\theta) = \{C' \mid ||f(\theta) - C'||_{\text{op}} \leq \widehat{q}\}$ *satisfies* $\mathcal{P}_{\Theta,\widetilde{C},\widehat{q}}(\widetilde{C} \in \mathcal{U}(\Theta)) \geq 1-\alpha$*, where* $||\cdot||_{\text{op}}$ *denotes the matrix operator norm. If for any* $\theta \in \Theta$ *and* $\delta > 0$*,* $\mathcal{P}(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta) > \mathcal{P}(\widehat{q}^2 \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta)$*, then*

$$\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{U}(\Theta)) \geq \mathcal{P}_{\Theta,\widetilde{C},\widehat{q}}(\widetilde{C} \in \mathcal{U}(\Theta)) \geq 1-\alpha.$$

3.5 Optimization Algorithm

Due to our generalization over traditional approaches to robust control (discussed in detail in Section 4), the standard approaches of solution used in those cases, namely generalized algebraic Riccati equations (GARE) or policy gradient, cannot be applied without modification. We, thus, now discuss how policy gradient can be adapted to efficiently solve the problem of interest and then demonstrate corresponding convergence results in Section 3.6. Given the novelty of the framing of Equation (3) over previous framings, specifically in the geometry of the uncertainty regions, the policy gradient expressions derived here too are novel, highlighted below. We frame this discussion around the *deterministic*, discrete-time, infinite time horizon setting. We assume that the initial state is drawn from a known distribution $x(0) \sim \mathcal{N}(0, X_0)$. Naively, computing the gradient would require estimation of the infinite sum in J; however, it is well known that the gradient can be computed using a Lyapunov formulation (Fazel et al., 2018), given by

$$\nabla_K J(K,A,B) = 2((R + B^\top P_K B)K - B^\top P_K A)X_K, \tag{3}$$

where X_K and P_K respectively solve the two Lyapunov equations $\ell_X(X_K, \Delta_K) = 0$ and $\ell_P(P_K, \Delta_K, K) = 0$ for specified Q, R, K, and $\Delta_K := A - BK$, where

$$\begin{aligned} \ell_X(X_K,\Delta_K) &:= \Delta_K X_K \Delta_K^\top - X_K + X_0 \\ \ell_P(P_K,\Delta_K,K) &:= \Delta_K^\top P_K \Delta_K + Q + K^\top R K - P_K \end{aligned} \tag{4}$$

Note that, while ℓ_P also depends on the choice of Q and R, we do not explicitly note this in the notation as they remain fixed throughout the problem. If the continuous-time setting is of interest instead, there are analogous Lyapunov equations and gradient expressions to those respectively in Equation (3) and Equation (4). To solve Equation (3), we wish to perform gradient updates on K instead with respect to $\phi(K) := \max_{\widehat{C} \in \mathcal{U}(\theta)} J(K,\widehat{C})$. Naively, one could proceed through the remaining analysis by leveraging Danskin's Theorem to compute the gradient of $\nabla_K \phi(K)$, which would result in the expression $\nabla_K \phi(K) = \nabla_K J(K, C^*(K))$, where $C^*(K) := \arg\max_{\widehat{C} \in \mathcal{U}(\theta)} J(K,\widehat{C})$; however, the existence of such

a gradient requires that $C^*(K)$ be the unique maximizer. Such an assumption is unlikely to hold in practice; for this reason, we instead relax this assumption and proceed using subgradients. That is, we suppose $C^*(K) := \text{Arg max}_{\widehat{C} \in \mathcal{U}(\theta)} J(K, \widehat{C})$ instead is a *set* and denote by $\partial_K \phi(K) := \{\nabla_K J(K, C_K) : C_K \in C^*(K)\}$ the *vertices* of the subdifferential.

Thus, robust policy optimization proceeds by iteratively updating K with *any* vertex of the subdifferential, namely by evaluating Equation (3) with $[A_K, B_K] := C_K$ for some $C_K \in C^*(K)$ and (X_K^*, P_K^*), the solutions to the Lyapunov equations when $\Delta_K^* := A_K - B_K K$. We initialize this procedure with the optimal controller for the nominally predicted dynamics, i.e. $K^{(0)} := K^*(f(\theta))$. Extending LQR policy gradient methods to the robust setting, therefore, reduces to being able to efficiently solve the maximization problem of C_K over $\mathcal{U}(\theta) := \mathcal{B}_{\widehat{q}}(f(\theta))$. This can be estimated with gradient ascent, where the Lyapunov expression $\nabla_C J(K, C) = 2P_K C W X_K W^\top$ is derived in Appendix H. The use of subgradients for policy optimization and the derivation of the explicit $\nabla_C J(K, C)$ expression in its Lyapunov formulation are novel contributions to the robust LQR space; these aspects were heretofore unstudied as previously studied robust formulations (in Section 4) could be translated into GAREs and, therefore, did not require algorithmic innovation. The algorithm is given in Algorithm 1.

Algorithm 1 CONFORMALIZED PREDICT-THEN-CONTROL (CPC)

procedure CPC($\theta, f(\theta), \widehat{q}, \eta_K, \eta_C, T_K, T_C$)
Inputs: Design θ, Predictor $f(\theta)$, Conformal quantile $\widehat{q}$, Step sizes η_K, η_C, Max steps T_K, T_C
  $\widehat{C} := f(\theta), K^{(0)} \leftarrow \text{SOLVEARE}(\widehat{C})$
  for $t_K \in \{0, \dots T_K - 1\}$ **do**
    $[A^{(0)}, B^{(0)}] := C^{(0)} \leftarrow \widehat{C}$
    for $t_C \in \{0, \dots T_C - 1\}$ **do**
      $\Delta^{(t_C)} := A^{(t_C)} - B^{(t_C)} K^{(t_K)}$
      $X^{(t_C)} \leftarrow \text{Solve}(\ell_X(X, \Delta^{(t_C)}) = 0; X)$
      $P^{(t_C)} \leftarrow \text{Solve}(\ell_P(P, \Delta^{(t_C)}, K^{(t_K)}) = 0; P)$
      $C^{(t_C+1)} \leftarrow \Pi_{\mathcal{B}_{\widehat{q}}(\widehat{C})}\big(C^{(t_C)} + \eta_C \left(2P^{(t_C)} C^{(t_C)} W^{(t_K)} X^{(t_C)} (W^{(t_K)})^\top\right)\big)$
    $[A_K, B_K] \leftarrow C^{(T_C)}$
    $\Delta^* \leftarrow A_K - B_K K^{(t_K)}$
    $X^* \leftarrow \text{Solve}(\ell_X(X, \Delta^*) = 0; X)$
    $P^* \leftarrow \text{Solve}(\ell_P(P, \Delta^*, K^{(t_K)}) = 0; P)$
    $K^{(t_K+1)} \leftarrow K^{(t_K)} - \eta_K\big(2((R + (B_K)^\top P^* B_K) K^{(t_K)} - (B_K)^\top P^* A_K) X^*\big)$
  Return $K^{(T_K)}$
end procedure

3.6 Policy Gradient Convergence Guarantees

We now wish to demonstrate this policy gradient approach retains the desired convergence properties it satisfies in the nominal case. Convergence guarantees surprisingly hold in the standard case despite the nonconvexity of the problem in K due to a property known as "gradient dominance" (Gravell, Esfahani, and Summers, 2020a). A function $f : \mathbb{R}^{d_1 \times d_2} \to \mathbb{R}$ is gradient-dominated if, for some $\mu > 0$, $f(x) - f(x^*) \leq \mu ||\nabla_x f(x)||_F^2$, where $x^* := \arg\min_x f(x)$.

We proceed through the analysis similarly leveraging gradient dominance; however, our analysis has the novel problem of having to handle the non-uniqueness of the subgradient being used, namely that our algorithm may perform updates with one of the collection of subdifferential vertices rather than using the uniquely defined gradient. For this reason, we instead consider a generalized notion of *subgradient* domination, defined as $\exists$ some $\mu > 0$ such that $f(x) - f(x^*) \leq \mu \min_{g \in \partial f(x)} ||g||_F^2$. We show that ϕ satisfies subgradient dominance and that this then produces convergence guarantees for Algorithm 1 in Lemma I.3.

The full proof is deferred to Appendix I and parallels the proof strategy presented in (Fazel et al., 2018); the main technical challenges are in demonstrating that bounds on expressions related to $J(K, C)$ and $\nabla_K J(K, C)$ are retained in our robust setting and that the non-uniqueness of the maximizer does not interfere with convergence. Intuitively, the non-uniqueness manifests as a looser gradient dominance constant and, thus, convergence decay rate, since μ must be taken to be the loosest constant amongst those of the maximizing set. In line with (Fazel et al., 2018), we assume $X_K \succcurlyeq 0$ across $\widehat{C} \in \mathcal{C}$ and $K \in \mathcal{K}(\mathcal{C})$. This is true if the system is controllable for any $\widehat{C} \in \mathcal{C}$, which holds if the nominal dynamics are well-behaved and the predictor $f(\theta)$ is sufficiently accurate, resulting in a small $\mathcal{C}$ set. The statements below are made for a general set of dynamics $\mathcal{C}$, though we are interested in $\mathcal{C} := \mathcal{U}(\theta)$. We defer the presentation of the explicit poly-expression in Theorem 3.8 to Appendix I.

Theorem 3.8. *Let* $J(K, C) := \sum_{t=0}^{\infty} (x_t^\top (Q_t + K^\top R_t K) x_t)$ *for* $w = 0$. *Let* $K^{(t)}$, $\phi(K) := \max_{C \in \mathcal{C}} J(K, C)$, *and* $K_{\text{rob}}^*(\mathcal{C}) := \arg\min_{K \in \mathcal{K}(\mathcal{C})} \phi(K)$ *be the t-th iterate of Algorithm 1. Assume for each iterate t, the optimization over C converges, i.e.* $C^{(T_C)} = C^*(K^{(t)})$, *that* $K^{(t)} \in \mathcal{K}(\mathcal{C})$, *and that* $X_K \succcurlyeq 0$ *for all* $\widehat{C} \in \mathcal{C}$ *and* $K \in \mathcal{K}(\mathcal{C})$. *Denote* $\nu := \min_{\widehat{C} \in \mathcal{C}} \min_{K \in \mathcal{K}(\mathcal{C})} \sigma_{\min}(X_K)$. *If in Algorithm 1* $\eta_K \leq \min_{[\widehat{A}, \widehat{B}] \in \mathcal{C}} \text{poly}(\frac{\nu \sigma_{\min}(Q)}{J(K^{(0)}, C)}, \frac{1}{||\widehat{A}||}, \frac{1}{||\widehat{B}||}, \frac{1}{||R||}, \sigma_{\min}(R))$, *then, there exists a* $\gamma > 0$ *such that* $\phi(K^{(T)}) - \phi(K_{\text{rob}}^*(\mathcal{C})) \leq (1-\gamma)^T (\phi(K_0) - \phi(K_{\text{rob}}^*(\mathcal{C})))$.

Formally, such convergence is guaranteed only if iterates $K^{(t)}$ remain within $\mathcal{K}(\mathcal{U}(\theta))$. One modification to Algorithm 1 would involve projecting intermediate iterates to this stabilizing set by solving

$$\begin{aligned} \Pi_{\mathcal{K}(\mathcal{U}(\theta))}(\widetilde{K}^{(t)}) := \arg\min_{K} ||K - \widetilde{K}^{(t)}||_{\text{op}} \\ \text{s.t.} \quad \max_{[\widehat{A},\widehat{B}]:=\widehat{C}\in\mathcal{U}(\theta)} \rho(\widehat{A} - \widehat{B}K) < 1. \end{aligned} \tag{5}$$

There, however, is no known efficient algorithm to solve this projection step. Despite being of theoretical concern, this instability issue fails to be practically relevant, since the controller iterates remain well within the set of stabilization for sufficiently accurate predictors $f(\theta)$. If instabilities arise, an approximate solution can be obtained by replacing $\max_{\widehat{C}\in\mathcal{U}(\theta)}$ of Equation (5) with a finite sampling $\{\widehat{C}^{(i)}\}$ over $\mathcal{U}(\theta)$.

4 RELATED WORKS

Robust control can be broadly categorized into trajectory-based and trajectory-free robustness. The former adjusts an initially posited control scheme in an *online* fashion based on feedback measurements (Azad and Herber, 2022; Seiler, Packard, and Gahinet, 2020; Paraskevopoulos, 2017), whereas the latter directly incorporates desired robustness into the optimization problem *prior* to deployment (Gravell and Summers, 2020; Gravell, Shames, and Summers, 2022). Given that control co-design seeks to identify a controller *prior* to deploying a design, we specifically highlight methods of trajectory-free robust control.

A popular classical trajectory-free method is $\mathcal{H}_\infty$ control. $\mathcal{H}_\infty$ is typically formulated as minimizing $||T_{wz}||_\infty$, i.e. the frequency space transfer function from $w \to z$ for some performance state z. By defining $z = \begin{bmatrix} Q^{1/2}x & R^{1/2}u \end{bmatrix}$, we recover a recognizable LQR formulation, with the objective replaced by

$$u^* = \min_{\{u_t\}} \max_{\{w_t\}} \sum_{t=0}^{\infty} (x_t^\top Q x_t + u_t^\top R u_t - \gamma^2 w_t^\top w_t), \tag{6}$$

which can be solved via generalized Riccati equations (Başar and Bernhard, 2008). Here, γ can either be fixed to perform suboptimal $\mathcal{H}_\infty$ synthesis or it can be determined via bisection to identify the smallest γ such that a solution exists. Notably, the nominal $\mathcal{H}_\infty$ formulation seeks additive, unstructured disturbance rejection. Of interest herein, however, was robustness to *multiplicative* uncertainties through the system dynamics. Towards this end, μ-synthesis offers an extension to $\mathcal{H}_\infty$ control by allowing users to specify norm-bounded uncertainties on system dynamics (Bevrani, Feizi, and Ataee, 2015; Chen, Yao, and Wang, 2014).

This need for manual specification in μ-synthesis, however, incurs conservatism or controller instability if poorly specified, resulting in increasing interest in data-driven specifications. In this vein, a formulation known as LQR with multiplicative noise (LQRm), has recently become of interest, where the controller is:

$$K^* := \arg\min_{K} \mathbb{E}_{\{\delta_i\},\{\gamma_i\}}[J(K, A, B)] \tag{7}$$

$$\dot{x} := \left(A + \sum_{i=1}^{p} \delta_i A_i\right) x + \left(B + \sum_{i=1}^{q} \gamma_i B_i\right) u + w,$$

where $\{A_i\}$ and $\{B_i\}$ and the distributions of $\{\delta_i\} \sim \mathcal{D}_\delta$ and $\{\gamma_i\} \sim \mathcal{D}_\gamma$ can be specified, either with data-free or with data-driven estimation. Most common among data-free specifications are so-called "margin methods." Briefly, margin methods specify $\{\delta_i\}$ and $\{\gamma_i\}$ by finding those $\{\delta_i\}$ and $\{\gamma_i\}$ that result in borderline-stable dynamics when paired with the corresponding, manually specified $(\{A_i\}, \{B_i\})$ and some choice of controller: the particular controller varies across margin strategies. A full description of the margin methods considered is given in Appendix J.

As with $\mathcal{H}_\infty$, such data-free LQRm methods sacrifice stability or risk conservatism in ignoring the nature of the dynamics predictor misspecification, resulting in recent works that give data-driven parameterizations (Gravell, Esfahani, and Summers, 2020b). Here, two approaches were proposed to learn the margin parameters, which we refer to as the "Shared Lyapunov" and "Auxiliary Stabilizer" approaches, described fully in their paper. While such approaches improve upon the conservatism of classical data-free margin methods, they still require the hand specification of the perturbation matrices $\{A_i\}$ and $\{B_i\}$.

5 EXPERIMENTS

We now study five setups of interest in the infinite horizon, discrete-time, deterministic setting, namely LQR control of an airfoil (Chrif and Kadda, 2014), a load positioning system (Ahmadi, Rahmani, and Shahmansoorian, 2023; Jiang et al., 2016), a Furuta pendulum (Arulmozhi and Victorie, 2022), a DC microgrid (Liu et al., 2023), and a fusion plant (Kirgni and Wang, 2023). The dimensions of (θ, A, B) are $(\mathbb{R}^{15}, \mathbb{R}^{4\times4}, \mathbb{R}^{4\times2})$ for the airfoil, $(\mathbb{R}^{5}, \mathbb{R}^{4\times4}, \mathbb{R}^{4\times1})$ for the load positioning system, $(\mathbb{R}^{9}, \mathbb{R}^{4\times4}, \mathbb{R}^{4\times1})$ for the Furuta pendulum, $(\mathbb{R}^{17}, \mathbb{R}^{9\times9}, \mathbb{R}^{9\times1})$ for the DC microgrid, and $(\mathbb{R}^{26}, \mathbb{R}^{8\times8}, \mathbb{R}^{8\times1})$ for the fusion plant. The full setup details are provided in Appendix K.

We compare against $\mathcal{H}_\infty$ control with γ bisection, the data-free margin methods, and the data-based methods for the LQRm setup as discussed in Section 4. The data-free margin methods are as implemented by

Table 1: Each of the results below are the p-values of paired t-tests conducted pairwise between methods testing $H_1 : \mathcal{R}_{\%}^{(\text{CPC})} < \mathcal{R}_{\%}^{(\text{alt})}$ over 1,000 i.i.d. test samples. For any comparison method with $> 80\%$ unstable cases (see Table 8 for percentages), we have marked the entry with "—".

	Airfoil	Load Positioning	Furuta Pendulum	DC Microgrids	Fusion Plant
Random Critical	—	—	—	—	—
Random OL MSS (Weak)	**0.0003**	—	—	—	—
Random OL MSUS	—	—	—	—	—
Row-Col Critical	—	—	—	—	—
Row-Col OL MSS (Weak)	**0.0117**	—	—	—	—
Row-Col OL MSUS	**0.0009**	—	—	—	—
Shared Lyapunov	**0.0001**	**0.0000**	**0.0112**	0.0913	**0.0023**
Auxiliary Stabilizer	**0.0001**	**0.0005**	**0.0055**	**0.0428**	0.0630
$\mathcal{H}_\infty$	**0.0009**	**0.0000**	**0.0071**	**0.0004**	**0.0013**

(Gravell and Summers, 2020) and are fully described in Appendix J, of which we specifically consider "Random Critical," "Random OL MSS (Weak)," "Random OL MSUS," "Row-Col Critical," "Row-Col OL MSS (Weak)," and "Row-Col OL MSUS." The data-based methods are the "Shared Lyapunov" and "Auxiliary Stabilizer" approaches from (Gravell, Esfahani, and Summers, 2020b). As discussed, we are considering the trajectory-*free* setting, so we do not compare against methods that achieve robustness adaptively over trajectories, such as those in (Gravell and Summers, 2020; Gravell, Shames, and Summers, 2022).

In the experiments, we construct $\mathcal{D}$ as per Section 3.1, using random gain matrices $K^{(i)}$. N was taken to be $2,000$ with $|\mathcal{D}_{\mathcal{C}}| = 400$ and the remaining $\mathcal{D}_T$ used to train $f(\theta)$, taken to be feed-forward neural networks.

5.1 Robust Control Regret & Stability

We first study the empirical regret across the aforementioned systems and robust control methods over $1,000$ i.i.d. test points from $\mathcal{P}(\Theta, C)$. To make results comparable across $\theta^{(i)}$, we normalize each trial by its nominal objective, i.e., $\mathcal{R}_{\%} = \mathcal{R}(\Theta, C)/J(K^*(C), C)$, as in (Sun, Liu, and Li, 2023). If the uncertainty regions of the robust problems are poorly specified, i.e. if the regions of robustness do not capture the true dynamics, the resulting robust controller may have unbounded cost, i.e. $J(K^*_{\text{rob}}, C) = \infty$. We, thus, only compute $\mathcal{R}_{\%}$ over the stabilizing controllers and separately report the proportion of destabilized cases. Lower values are desirable for both.

For each comparison method, we report the result of a one-side paired t-test of $H_1 : \mathcal{R}_{\%}^{(\text{CPC})} < \mathcal{R}_{\%}^{(\text{alt})}$ in Table 1. We defer the presentation of the raw regret values and the percent of cases with stabilized dynamics to Appendix L due to space constraints. Notably, the alternative approaches generally incur greater regret

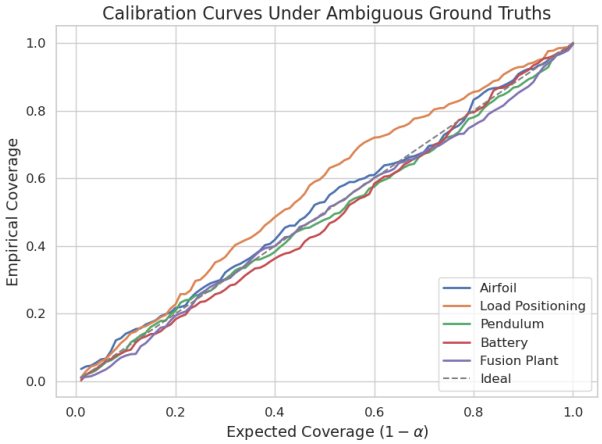

Figure 1: Calibration plots for the tasks, assessed on 1,000 i.i.d. test samples of C with calibration performed using the estimated $\widetilde{C}$, affirming Theorem 3.7.

than CPC. For $\mathcal{H}_\infty$, this is expected as the misspecification here is in the dynamics matrices, differing from the adversarial exogenous noise that $\mathcal{H}_\infty$ is designed to protect against. Similarly, the data-free margin methods protect against perturbations that are misaligned with the true dynamics misspecification, which result in significant instability for the higher-dimensional problems (i.e. the "Furuta Pendulum," "DC Microgrids," and "Fusion Plant" tasks). The data-driven LQRm methods improve significantly upon these margin approaches in stability, yet they are too conservative as they do not make use of the anticipated structures of the errors made in the predictions by $\widehat{f}(\theta)$.

5.2 Ambiguous Ground Truth Calibration

To validate the results of Theorem 3.7 and demonstrate the empirical validity of the associated assump-

tion, we computed the empirical coverages across various levels of desired coverage $\alpha \in (0, 1)$ for the experimental setups. As previously discussed, the calibration here was performed using a calibration set of *estimated* dynamics $\widetilde{\mathcal{D}} = \{(\theta^{(i)}, \widetilde{C}^{(i)})\}$ but coverage was assessed on the *true* dynamics $\{C^{(i)}\}$. We computed this in the manner described in Section 5 for α varying by increments of 0.05. For assessing coverage, we again used $1,000$ test points drawn i.i.d. from $\mathcal{P}(\Theta, C)$ and measured the proportion of samples for which $s(\theta^{(i)}, C^{(i)}) \leq \widehat{q}$. The results are shown in Figure 1, where we see the desired calibration under calibration with estimated dynamics.

6 DISCUSSION

We have presented CPC, a principled framework for specifying the LQR robust control subproblem in a UCCD setting, suggesting many directions for extension. The most immediate would involve integrating this framework fully into a UCCD pipeline: we focused herein on the robust control subproblem but characterizing the end-to-end workflow is of great interest. In addition, nonlinear extension by leveraging Koopman operator theory or nonparametric neural operator models would be interesting (Brunton et al., 2021; Mauroy, Susuki, and Mezić, 2020; Qian et al., 2020) as would the extension to MDPs (Wang et al., 2021).

References

[1] Peyman Ahmadi, Mehdi Rahmani, and Aref Shahmansoorian. "LQR based optimal co-design for linear control systems with input and state constraints". In: *International Journal of Systems Science* 54.5 (2023), pp. 1136–1149.

[2] Christopher T Aksland et al. "An Approach to Robust Co-Design of Plant and Closed-Loop Controller". In: *2023 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE. 2023, pp. 918–925.

[3] James T Allison, Tinghao Guo, and Zhi Han. "Co-design of an active suspension using simultaneous dynamic optimization". In: *Journal of Mechanical Design* 136.8 (2014), p. 081003.

[4] Anastasios N Angelopoulos and Stephen Bates. "A gentle introduction to conformal prediction and distribution-free uncertainty quantification". In: *arXiv preprint arXiv:2107.07511* (2021).

[5] N Arulmozhi and T Victorie. "Kalman Filter and H∞ Filter Based Linear Quadratic Regulator for Furuta Pendulum." In: *Computer Systems Science & Engineering* 43.2 (2022).

[6] Saeed Azad and Michael J Alexander-Ramos. "A single-loop reliability-based MDSDO formulation for combined design and control optimization of stochastic dynamic systems". In: *Journal of Mechanical Design* 143.2 (2020), p. 021703.

[7] Saeed Azad and Michael J Alexander-Ramos. "Robust combined design and control optimization of hybrid-electric vehicles using MDSDO". In: *IEEE Transactions on Vehicular Technology* 70.5 (2021), pp. 4139–4152.

[8] Saeed Azad and Michael J Alexander-Ramos. "Robust MDSDO for co-design of stochastic dynamic systems". In: *Journal of Mechanical design* 142.1 (2020), p. 011403.

[9] Saeed Azad and Daniel R Herber. "An overview of uncertain control co-design formulations". In: *Journal of Mechanical Design* 145.9 (2023), p. 091709.

[10] Saeed Azad and Daniel R Herber. "Concurrent Probabilistic Control Co-Design and Layout Optimization of Wave Energy Converter Farms Using Surrogate Modeling". In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 87318. American Society of Mechanical Engineers. 2023, V03BT03A035.

[11] Saeed Azad and Daniel R Herber. "Control co-design under uncertainties: formulations". In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 86229. American Society of Mechanical Engineers. 2022, V03AT03A008.

[12] Saeed Azad et al. "Comprehensive PHEV powertrain co-design performance studies using MDSDO". In: *Advances in Structural and Multidisciplinary Optimization: Proceedings of the 12th World Congress of Structural and Multidisciplinary Optimization (WCSMO12) 12*. Springer. 2018, pp. 83–97.

[13] Saeed Azad et al. "PHEV powertrain co-design with vehicle performance considerations using MDSDO". In: *Structural and Multidisciplinary Optimization* 60 (2019), pp. 1155–1169.

[14] Saeed Azad et al. "Site-dependent Solutions of Wave Energy Converter Farms with Surrogate Models, Control Co-design, and Layout Optimization". In: *arXiv preprint arXiv:2405.06794* (2024).

[15] Tamer Başar and Pierre Bernhard. *H-infinity optimal control and related minimax design problems: a dynamic game approach*. Springer Science & Business Media, 2008.

[16] Mohammad Behtash and Michael J Alexander-Ramos. "A Comparative Study Between the Generalized Polynomial Chaos Expansion-and First-Order Reliability Method-Based Formulations of Simulation-Based Control Co-Design". In: *Journal of Mechanical Design* (2024), pp. 1–17.

[17] Mohammad Behtash and Michael J Alexander-Ramos. "Decomposition-based MDSDO For co-design of large-scale dynamic systems". In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 51753. American Society of Mechanical Engineers. 2018, V02AT03A003.

[18] Petar Bevanda et al. "Towards Data-driven LQR with Koopmanizing Flows". In: *IFAC-PapersOnLine* 55.15 (2022), pp. 13–18.

[19] Hassan Bevrani, Mohammad Ramin Feizi, and Sirwan Ataee. "Robust frequency control in an islanded microgrid: H_∞ and μ-synthesis approaches". In: *IEEE transactions on smart grid* 7.2 (2015), pp. 706–717.

[20] Trevor J Bird et al. "A Set-Based Approach for Robust Control Co-Design". In: *arXiv preprint arXiv:2310.11658* (2023).

[21] Steven L Brunton et al. "Modern Koopman theory for dynamical systems". In: *arXiv preprint arXiv:2102.12086* (2021).

[22] Jingjing Bu et al. "LQR through the lens of first order methods: Discrete-time case". In: *arXiv preprint arXiv:1907.08921* (2019).

[23] Prasad Vilas Chanekar, Nikhil Chopra, and Shapour Azarm. "Co-design of linear systems using Generalized Benders Decomposition". In: *Automatica* 89 (2018), pp. 180–193.

[24] Zheng Chen, Bin Yao, and Qingfeng Wang. "μ-Synthesis-Based Adaptive Robust Control of Linear Motor Driven Stages With High-Frequency Dynamics: A Case Study". In: *IEEE/ASME Transactions on Mechatronics* 20.3 (2014), pp. 1482–1490.

[25] Abhilash Chenreddy and Erick Delage. "End-to-end Conditional Robust Optimization". In: *arXiv preprint arXiv:2403.04670* (2024).

[26] Abhilash Reddy Chenreddy, Nymisha Bandi, and Erick Delage. "Data-driven conditional robust optimization". In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 9525–9537.

[27] Labane Chrif and Zemalache Meguenni Kadda. "Aircraft control system using LQG and LQR controller with optimal estimation-Kalman filter design". In: *Procedia Engineering* 80 (2014), pp. 245–257.

[28] Tonghui Cui, James T Allison, and Pingfeng Wang. "A comparative study of formulations and algorithms for reliability-based co-design problems". In: *Journal of Mechanical Design* 142.3 (2020), p. 031104.

[29] Tonghui Cui, James T Allison, and Pingfeng Wang. "Reliability-based co-design of state-constrained stochastic dynamical systems". In: *AIAA Scitech 2020 Forum*. 2020, p. 0413.

[30] Tonghui Cui, Zhuoyuan Zheng, and Pingfeng Wang. "Control co-design of lithium-ion batteries for enhanced fast-charging and cycle life performances". In: *Journal of Electrochemical Energy Conversion and Storage* 19.3 (2022), p. 031001.

[31] Robert Falck et al. "dymos: A Python package for optimal control of multidisciplinary systems". In: *Journal of Open Source Software* 6.59 (2021), p. 2809.

[32] Hosam K Fathy et al. "Nested plant/controller optimization with application to combined passive/active automotive suspensions". In: *Proceedings of the 2003 American Control Conference, 2003.* Vol. 4. IEEE. 2003, pp. 3375–3380.

[33] Hosam K Fathy et al. "On the coupling between the plant and controller optimization problems". In: *Proceedings of the 2001 American Control Conference.(Cat. No. 01CH37148).* Vol. 3. IEEE. 2001, pp. 1864–1869.

[34] Maryam Fazel et al. "Global convergence of policy gradient methods for the linear quadratic regulator". In: *International conference on machine learning*. PMLR. 2018, pp. 1467–1476.

[35] Shai Feldman et al. "Conformal Prediction is Robust to Dispersive Label Noise". In: *Conformal and Probabilistic Prediction with Applications*. PMLR. 2023, pp. 624–626.

[36] Bernard Friedland. *Advanced control system design*. Prentice-Hall, Inc., 1995.

[37] Mario Garcia-Sanz. "Control Co-Design: an engineering game changer". In: *Advanced Control for Applications: Engineering and Industrial Systems* 1.1 (2019), e18.

[38] Benjamin Gravell, Peyman Mohajerin Esfahani, and Tyler Summers. "Learning optimal controllers for linear systems with multiplicative noise via policy gradient". In: *IEEE Transactions on Automatic Control* 66.11 (2020), pp. 5283–5298.

[39] Benjamin Gravell, Iman Shames, and Tyler Summers. "Robust Data-Driven Output Feedback Control via Bootstrapped Multiplicative Noise". In: *Learning for Dynamics and Control Conference*. PMLR. 2022, pp. 650–662.

[40] Benjamin Gravell and Tyler Summers. "Robust learning-based control via bootstrapped multiplicative noise". In: *Learning for Dynamics and Control*. PMLR. 2020, pp. 599–607.

[41] Benjamin J Gravell, Peyman Mohajerin Esfahani, and Tyler H Summers. "Robust control design for linear systems via multiplicative noise". In: *IFAC-PapersOnLine* 53.2 (2020), pp. 7392–7399.

[42] Daniel R Herber and James T Allison. "Nested and simultaneous solution strategies for general combined plant and control design problems". In: *Journal of Mechanical Design* 141.1 (2019), p. 011402.

[43] Yu Jiang et al. "An iterative approach to the optimal co-design of linear control systems". In: *International Journal of Control* 89.4 (2016), pp. 680–690.

[44] Chancellor Johnstone and Bruce Cox. "Conformal uncertainty sets for robust optimization". In: *Conformal and Probabilistic Prediction and Applications*. PMLR. 2021, pp. 72–90.

[45] Abdullah Kamadan, Gullu Kiziltas, and Volkan Patoglu. "Co-design strategies for optimal variable stiffness actuation". In: *IEEE/ASME Transactions on Mechatronics* 22.6 (2017), pp. 2768–2779.

[46] Taylor Killian, George Konidaris, and Finale Doshi-Velez. "Transfer learning across patient variations with hidden parameter markov decision processes". In: *arXiv preprint arXiv:1612.00475* (2016).

[47] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).

[48] Hamza Boubacar Kirgni and Junling Wang. "LQR-based adaptive TSMC for nuclear reactor in load following operation". In: *Progress in Nuclear Energy* 156 (2023), p. 104560.

[49] Yulin Liu et al. "A novel online learning-based linear quadratic regulator for vanadium redox flow battery in DC microgrids". In: *Journal of Power Sources* 587 (2023), p. 233672.

[50] Lennart Ljung et al. "Theory for the user". In: *System identification* (1987).

[51] Giorgos Mamakoukas et al. "Local Koopman operators for data-driven control of robotic systems". In: *Robotics: science and systems*. 2019.

[52] Alexandre Mauroy, Y Susuki, and I Mezić. *Koopman operator in systems and control*. Springer, 2020.

[53] Austin L Nash, Herschel C Pangborn, and Neera Jain. "Robust control co-design with receding-horizon mpc". In: *2021 American Control Conference (ACC)*. IEEE. 2021, pp. 373–379.

[54] Vu Linh Nguyen, Chin-Hsing Kuo, and Po Ting Lin. "Reliability-based analysis and optimization of the gravity balancing performance of spring-articulated serial robots with uncertainties". In: *Journal of Mechanisms and Robotics* 14.3 (2022), p. 031016.

[55] Paraskevas N Paraskevopoulos. *Modern control engineering*. CRC Press, 2017.

[56] Adam Paszke et al. "Pytorch: An imperative style, high-performance deep learning library". In: *Advances in Neural Information Processing Systems* 32 (2019).

[57] Yash P Patel, Sahana Rayan, and Ambuj Tewari. "Conformal contextual robust optimization". In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 2024, pp. 2485–2493.

[58] Elizabeth Qian et al. "Lift & learn: Physics-informed machine learning for large-scale nonlinear dynamical systems". In: *Physica D: Nonlinear Phenomena* 406 (2020), p. 132401.

[59] Julie A Reyer et al. "Comparison of combined embodiment design and control optimization strategies using optimality conditions". In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 80234. American Society of Mechanical Engineers. 2001, pp. 1023–1032.

[60] Utsav Sadana et al. "A survey of contextual optimization methods for decision-making under uncertainty". In: *European Journal of Operational Research* (2024).

[61] Simo Särkkä and Arno Solin. *Applied stochastic differential equations*. Vol. 10. Cambridge University Press, 2019.

[62] Peter Seiler, Andrew Packard, and Pascal Gahinet. "An introduction to disk margins [lecture notes]". In: *IEEE Control Systems Magazine* 40.5 (2020), pp. 78–95.

[63] Glenn Shafer and Vladimir Vovk. "A Tutorial on Conformal Prediction." In: *Journal of Machine Learning Research* 9.3 (2008).

[64] Chunlin Sun, Linyu Liu, and Xiaocheng Li. "Predict-then-calibrate: A new perspective of robust contextual lp". In: *Advances in neural information processing systems* 36 (2023), pp. 17713–17741.

[65] Yue Sun and Maryam Fazel. "Learning optimal controllers by policy gradient: Global optimality via convex parameterization". In: *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE. 2021, pp. 4576–4581.

[66] Russ Tedrake. *Underactuated Robotics. Algorithms for Walking, Running, Swimming, Flying, and Manipulation*. 2023.

[67] Kai Wang et al. "Learning mdps from features: Predict-then-optimize for sequential decision making by reinforcement learning". In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 8795–8806.

[68] Jan Willems. "Least squares stationary optimal control and the algebraic Riccati equation". In: *IEEE Transactions on automatic control* 16.6 (1971), pp. 621–634.

[69] Yi Wu et al. "Learning and planning with a semantic model". In: *arXiv preprint arXiv:1809.10842* (2018).

[70] Christopher Yeh et al. "End-to-end conformal calibration for optimization under uncertainty". In: *arXiv preprint arXiv:2409.20534* (2024).

[71] Jerzy Zabczyk. *Mathematical control theory*. Springer, 2020.

[72] Dongdong Zhao et al. "A Kalman–Koopman LQR Control Approach to Robotic Systems". In: *IEEE Transactions on Industrial Electronics* (2024).

A Prediction Region Validity Lemma

Given that we characterize *both* the continuous- and discrete-time settings below, we produce a generalized definition to that presented in Definition 3.4.

Definition A.1. Let $M(C, K^*(C)) := A - BK^*(C)$ be the optimal closed-loop matrix. Define

$$r(C, K^*(C)) := \begin{cases} \frac{\min_i -\mathrm{Re}(\lambda_i(M))}{2\kappa(U)\,||W||_{\mathrm{op}}} & \text{Continuous time setting} \\ \frac{\min_i (1-|\lambda_i(M)|)}{2\kappa(U)\,||W||_{\mathrm{op}}} & \text{Discrete time setting} \end{cases}$$

where $M = U\Lambda U^{-1}$, $\kappa(U)$ is the condition number of U, and $W = [I \ -K^*(C)^\top]^\top$.

Lemma A.2. *Let $J(K, C)$ be a function such that, for any fixed θ, it is non-negative and L-Lipschitz in $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ under the operator norm for any $K \in \mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$, where $\mathcal{K} : \Omega(\Theta) \to \Omega(\mathcal{C})$ and $\Omega_1 \subset \Omega_2 \implies \mathcal{K}(\Omega_2) \subset \mathcal{K}(\Omega_1)$. Further assume that $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\theta))) \geq 1 - \alpha$. Then:*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\left(0 \leq \mathcal{R}(\Theta, C) \leq 2L\widehat{q} + \Delta_{\mathrm{dom}}(\Theta, C)\right) \geq 1 - \alpha. \tag{8}$$

Further, if $\widehat{q} < r(C, K^*(C))$, (see Definition A.1) $\Delta_{\mathrm{dom}}(\Theta, C) = 0$.

Proof. We consider the event of interest conditionally on a pair (θ, C) where $C \in \mathcal{B}_{\widehat{q}}(f(\theta))$. By assumption, we then have that $\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta))) \subset \mathcal{K}(C)$. As previously noted, the suboptimality here is defined over the *true* C matrix, meaning, unlike previous works, we here wish to bound $J(K^*(\mathcal{B}_{\widehat{q}}(f(\theta))), C) - \min_{K\in\mathcal{K}(C)} J(K, C)$ in place of $\min_{K\in\mathcal{K}(C)} \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} J(K, \widehat{C}) - \min_{K\in\mathcal{K}(C)} J(K, C)$, where $K^*(\mathcal{B}_{\widehat{q}}(f(\theta))) := \arg\min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} J(K, \widehat{C})$. We begin by matching the minimization sets in the terms as follows:

$$\begin{aligned} &\left| J(K^*(\mathcal{B}_{\widehat{q}}(f(\theta))), C) - \min_{K\in\mathcal{K}(C)} J(K, C) \right| \\ &= \left| J(K^*(\mathcal{B}_{\widehat{q}}(f(\theta))), C) - \min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} J(K, C) + \min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} J(K, C) - \min_{K\in\mathcal{K}(C)} J(K, C) \right| \\ &\leq \underbrace{\left| J(K^*(\mathcal{B}_{\widehat{q}}(f(\theta))), C) - \min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} J(K, C) \right|}_{\text{statistical robustness cost}} + \underbrace{\left| \min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} J(K, C) - \min_{K\in\mathcal{K}(C)} J(K, C) \right|}_{\text{universal stabilization cost}} \end{aligned}$$

As discussed in the main text, the error decomposes into two terms: the first from making the controller robust to adversarial dynamics matrices and the second from requiring that such a controller stabilize the whole collection of dynamics. We now bound each term separately, starting with the first term. We first note:

$$J(K^*(\mathcal{B}_{\widehat{q}}(f(\theta))), C) \leq \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} J(K^*(\mathcal{B}_{\widehat{q}}(f(\theta))), \widehat{C}) =: \min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} J(K, \widehat{C})$$

where the first step follows by the assumption $C \in \mathcal{B}_{\widehat{q}}(f(\theta))$ and second by definition of $K^*(\mathcal{B}_{\widehat{q}}(f(\theta)))$. From here,

$$\begin{aligned} \left| J(K^*(\mathcal{B}_{\widehat{q}}(f(\theta))), C) - \min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} J(K, C) \right| &\leq \left| \min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} J(K, \widehat{C}) - \min_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} J(K, C) \right| \\ &\leq \max_{K\in\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))} \left| \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} J(K, \widehat{C}) - J(K, C) \right| \\ &\leq L \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} ||\widehat{C} - C||_{\mathrm{op}} \leq 2L\widehat{q}. \end{aligned}$$

We now demonstrate that the second term vanishes within a radius of "safety," which we do through perturbation analysis of the closed-loop margin. In particular, notice that, since $\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta))) \subset \mathcal{K}(C)$, this term vanishes if the minimizer over $\mathcal{K}(C)$ lies in $\mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$. That is, this difference vanishes if $K^*(C) := \arg\min_{K\in\mathcal{K}(C)} J(K, C)$ satisfies $K^*(C) \in \mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$. Notably, this is equivalent to finding a condition under which $K^*(C)$ stabilizes all the dynamics matrices $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$.

To procure a test of this property, we consider an approach from the theory of switched linear systems, where controllers are sought that stabilize a collection of adjusting dynamics matrices. In particular, recall that $\mathcal{B}_{\widehat{q}}(f(\theta))$ is constructed under an operator norm and, therefore, any $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ can be viewed as a bounded perturbation to the nominal prediction corresponding to θ. That is, that $\widehat{C} = C + \Delta$ for $||\Delta||_{\text{op}} \leq 2\widehat{q}$. Thus, we can equivalently view the task as seeking a condition that guarantees that, if $||C - \widehat{C}||_{\text{op}} \leq 2\widehat{q}$, $K^*(C)$ stabilizes $\widehat{C}$.

We now consider the $W := [I_{n\times n} \; -K^*(C)^\top]^\top$ matrix and analyze the closed-loop stability of $\widehat{C}W$. By definition, we have that $\widehat{C}W := (C + \Delta)W =: CW + E$, where we know that CW is stabilized by $K^*(C)$, since $K^*(C) \in \mathcal{K}(C)$. By this latter point, we know CW satisfies one of two properties, depending on whether the system being analyzed is a continuous- or discrete-time setting. In the case of continuous time, we have that $\min_i -\text{Re}(\lambda_i(CW)) > 0$ and that, for discrete time, $\min_i(1 - |\lambda_i(CW)|) > 0$.

Under the additional assumption of CW being diagonalizable, we have that $CW = U\Lambda U^{-1}$ for $\Lambda = \text{diag}(\lambda_1, ..., \lambda_n)$. With this, we now return to analyzing the perturbed $CW + E$. By the Bauer–Fike bound, we know that for any eigenvalue λ'_j of $CW + E$, $\min_i |\lambda'_j - \lambda_i| \leq \kappa(U)||E||_{\text{op}}$, where $\kappa(U)$ is the condition number of U. Thus, if these perturbed eigenvalues remain within the stabilized regions, $\widehat{C}$ is stabilized by $K^*(C)$.

In the continuous-time setting, this is guaranteed if $\kappa(U)||E||_{\text{op}} < \min_i -\text{Re}(\lambda_i(CW))$ or, by the fact that $||E|| := ||\Delta W|| \leq ||\Delta|| \cdot ||W|| \leq 2\widehat{q}||W||$, if

$$\kappa(U)(2\widehat{q}||W||) < \min_i -\text{Re}(\lambda_i(CW)) \iff \widehat{q} < \frac{\min_i -\text{Re}(\lambda_i(CW))}{2\kappa(U)||W||}$$

The discrete-time setting follows analogously, simply with the stability condition replaced by the discrete-time analog, i.e. if $\kappa(U)||E||_{\text{op}} < \min_i(1 - |\lambda_i(CW)|)$. That is, a sufficient condition for stabilization is that

$$\kappa(U)(2\widehat{q}||W||) < \min_i(1 - |\lambda_i(CW)|) \iff \widehat{q} < \frac{\min_i(1 - |\lambda_i(CW)|)}{2\kappa(U)||W||}$$

Since we have the assumption that $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\theta))) \geq 1 - \alpha$, the result immediately follows. □

B Deterministic Discrete-Time Regret Analysis

Theorem B.1. *[Deterministic, discrete-time] Let $J(K, C) := \sum_{t=0}^{\infty}(x_t^\top(Q_t + K^\top R_t K)x_t)$ with $w = 0$. Assume that $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\Theta))) \geq 1 - \alpha$. Then, under Assumption 3.1,*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\,(0 \leq \mathcal{R}(\Theta, C) \leq 2L\widehat{q} + \Delta_{\text{dom}}(\Theta, C)) \geq 1 - \alpha,$$

where L is the Lipschitz constant of $J(K, \widehat{C})$ in $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ under the operator norm. Further, if $\widehat{q} < r(C, K^*(C))$, (see discrete-time in Definition A.1) $\Delta_{\text{dom}}(\Theta, C) = 0$.

Proof. We consider any fixed θ and demonstrate that $J(K, \widehat{C})$ is non-negative and L-Lipschitz in $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ under the operator norm for any $K \in \mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$, from which Lemma A.2 can be invoked to arrive at the desired conclusion. Given the assumed determinism of the dynamics, we have that $x_t = (\widehat{C}W)^t x_0$, meaning the above objective setup can equivalently be expressed as:

$$\sum_{t=0}^{\infty} x_0^\top((\widehat{C}W)^{t\top}(Q_t + K^\top R_t K)(\widehat{C}W)^t)x_0, \tag{9}$$

J is clearly non-negative by construction. It, therefore, suffices to demonstrate this objective is Lipschitz continuous with an appropriate Lipschitz constant. Notice the Lipschitz constant can be obtained by bounding the magnitude of the gradient with respect to $\widehat{C}$, which we do as follows

$$\begin{aligned}\nabla_{\widehat{C}}\left(\sum_{t=0}^{\infty} x_0^\top((\widehat{C}W)^{t\top}(Q_t + K^\top R_t K)(\widehat{C}W)^t)x_0\right) &= \sum_{t=0}^{\infty} t\text{diag}((Q_t + K^\top R_t K)(\widehat{C}W)^t x_0)(\widehat{C}W)^{(t-1)}\text{diag}(x_0)W^\top \\ &+ \sum_{t=0}^{\infty} t\text{diag}((Q_t^\top + (R_t K)^\top K)(\widehat{C}W)^t x_0)(\widehat{C}W)^{(t-1)}\text{diag}(x_0)W^\top\end{aligned}$$

We now bound the magnitude of this quantity as follows:

$$\begin{aligned}
L &\leq \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} || \sum_{t=0}^{\infty} t\text{diag}((Q_t + K^\top R_t K)(\widehat{C}W)^t x_0)(\widehat{C}W)^{(t-1)}\text{diag}(x_0)W^\top \\
&\qquad + \sum_{t=0}^{\infty} t\text{diag}((Q_t^\top + (R_t K)^\top K)(\widehat{C}W)^t x_0)(\widehat{C}W)^{(t-1)}\text{diag}(x_0)W^\top ||_{\text{op}} \\
&\leq \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t \left|\left|\text{diag}((Q_t + K^\top R_t K)(\widehat{C}W)^t x_0)(\widehat{C}W)^{(t-1)}\text{diag}(x_0)W^\top\right|\right|_{\text{op}} \\
&\qquad + \sum_{t=0}^{\infty} t \left|\left|\text{diag}((Q_t^\top + (R_t K)^\top K)(\widehat{C}W)^t x_0)(\widehat{C}W)^{(t-1)}\text{diag}(x_0)W^\top\right|\right|_{\text{op}},
\end{aligned}$$

where we have used $\text{diag}(x_0)$ for a vector x_0 to denote a diagonal matrix with x_0 placed along its main diagonal. We now bound each of these two terms separately, although the structure of the two is the same, so we explicitly show steps for bounding the first, from which the same can be repeated on the second. Importantly, we make use of the fact $||\text{diag}(x_0)||_{\text{op}} = ||x_0||_\infty$ and $||A||_\infty \leq \sqrt{n}||A||_{\text{op}}$ for $A \in \mathbb{R}^{n\times n}$ as follows:

$$\begin{aligned}
&\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t||\text{diag}((Q_t + K^\top R_t K)(\widehat{C}W)^t x_0)(\widehat{C}W)^{(t-1)}\text{diag}(x_0)W^\top||_{\text{op}} \\
&\qquad \leq \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t||\text{diag}((Q_t + K^\top R_t K)(\widehat{C}W)^t x_0)||_{\text{op}}||(\widehat{C}W)^{(t-1)}||_{\text{op}}||\text{diag}(x_0)||_{\text{op}}||W^\top||_{\text{op}} \\
&\qquad = \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t||(Q_t + K^\top R_t K)(\widehat{C}W)^t x_0||_\infty||x_0||_\infty||(\widehat{C}W)^{(t-1)}||_{\text{op}}||W||_{\text{op}} \\
&\qquad \leq \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t||Q_t + K^\top R_t K||_\infty||(\widehat{C}W)^t||_\infty||x_0||_\infty||x_0||_\infty||(\widehat{C}W)^{(t-1)}||_{\text{op}}||W||_{\text{op}} \\
&\qquad \leq \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t(\sqrt{n}||Q_t + K^\top R_t K||_\infty||x_0||_\infty^2||W||_{\text{op}})||(\widehat{C}W)^t||_{\text{op}}||(\widehat{C}W)^{(t-1)}||_{\text{op}}.
\end{aligned}$$

We now collect all terms independent of t into $D(K) = \max_{t\geq 0}\sqrt{n}||Q_t + K^\top R_t K||_\infty||x_0||_\infty^2||W||_{\text{op}}$:

$$\leq D(K) \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t||(\widehat{C}W)^t||_{\text{op}}||(\widehat{C}W)^{(t-1)}||_{\text{op}}.$$

Critically, we can now demonstrate that this sum is bounded by virtue of K being a universal stabilizer of the dynamics set $\mathcal{B}_{\widehat{q}}(f(\theta))$. By this stabilization, we know that $\widehat{C}W$ is Schur stable, i.e. $\min_i(1 - |\lambda_i(\widehat{C}W)|) > 0$ or $\max_i |\lambda_i(\widehat{C}W)| < 1$. Thus, there exists a $P \succ 0$ such that for some $\tau \in (0, 1)$, we have

$$(\widehat{C}W)^\top P(\widehat{C}W) \preceq \tau^2 P.$$

We now denote the norm induced by such a P as $||\cdot||_P$. Then, $||\widehat{C}W||_P \leq \tau$, meaning $||(\widehat{C}W)^t|| \leq \tau^t$. By the norm equivalence between the induced matrix norm and the standard operator norm, we have that

$$||(\widehat{C}W)^t||_{\text{op}} \leq \kappa(P)||(\widehat{C}W)^t||_P \leq \kappa(P)\tau^t,$$

where $\kappa(P) := \sqrt{\lambda_{\max}(P)/\lambda_{\min}(P)}$ is the condition number of P. We now see that the previous sum converges:

$$\begin{aligned}
D(K) \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t||(\widehat{C}W)^t||_{\text{op}}||(\widehat{C}W)^{(t-1)}||_{\text{op}} &\leq D(K)\kappa(P)^2 \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \sum_{t=0}^{\infty} t\tau^{2t-1} \\
&\leq D(K)\kappa(P)^2 \frac{\tau}{(1-\tau^2)^2},
\end{aligned}$$

completing the proof. □

The finite LQR case follows immediately as a corollary of the above, stated below for completeness.

Corollary B.2. *[Deterministic, discrete-time, finite-horizon] Let* $J(K,C) := \sum_{t=0}^{T}(x_t^\top(Q_t + K^\top R_t K)x_t)$ *with* $w = 0$. *Assume that* $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\Theta))) \geq 1-\alpha$. *Then:*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\left(0 \leq \mathcal{R}(\Theta,C) \leq 2L\widehat{q} + \Delta_{\text{dom}}(\Theta,C)\right) \geq 1-\alpha,$$

where L *is the Lipschitz constant of* $J(K,\widehat{C})$ *in* $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ *under the operator norm. Further, if* $\widehat{q} < r(C,K^*(C))$, *(see discrete-time in Definition A.1)* $\Delta_{\text{dom}}(\Theta,C) = 0$.

C Deterministic Continuous-Time Regret Analysis

The proof follows in much the same manner as the discrete-time case, with modest adjustments to the exact specification of the assumptions regarding the dynamics.

Theorem C.1. *[Deterministic, continuous-time] Let* $J(K,C) := \int_0^\infty (x(t)^\top(Q(t)+K^\top R(t)K)x(t))dt$ *for* $w = 0$. *Assume that* $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\Theta))) \geq 1-\alpha$. *Then, under Assumption 3.1,*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\left(0 \leq \mathcal{R}(\Theta,C) \leq 2L\widehat{q} + \Delta_{\text{dom}}(\Theta,C)\right) \geq 1-\alpha,$$

where L *is the Lipschitz constant of* $J(K,\widehat{C})$ *in* $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ *under the operator norm. Further, if* $\widehat{q} < r(C,K^*(C))$, *(see continuous-time in Definition A.1)* $\Delta_{\text{dom}}(\Theta,C) = 0$.

Proof. We consider any fixed θ and demonstrate the desired properties that $J(K,C)$ is non-negative and L-Lipschitz in $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ under the operator norm for any $K \in \mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$, from which Lemma A.2 can be invoked to arrive at the desired conclusion. Given the assumed determinism of the dynamics, we further have that $x(t) = e^{\widehat{C}Wt}x(0)$, meaning the above objective setup can equivalently be expressed with the uncertainty sets related to the objective function, namely as:

$$\int_0^\infty x(0)^\top (e^{\widehat{C}Wt})^\top (Q(t)+K^\top R(t)K)(e^{\widehat{C}Wt})x(0)dt. \tag{10}$$

J is clearly non-negative by construction. It, therefore, suffices to demonstrate this objective is Lipschitz continuous with an appropriate Lipschitz constant. We again proceed by bounding the norm of the gradient:

$$\begin{aligned}
&\nabla_{\widehat{C}}\left(\int_0^\infty x(0)^\top (e^{\widehat{C}Wt})^\top (Q(t)+K^\top R(t)K)(e^{\widehat{C}Wt})x(0)dt\right) \\
&= \int_0^\infty t\,\text{diag}((Q(t)+K^\top R(t)K)e^{\widehat{C}Wt}x(0))e^{\widehat{C}Wt}\text{diag}(x(0))W^\top dt \\
&+ \int_0^\infty t\,\text{diag}((Q(t)^\top + (R(t)K)^\top K)e^{\widehat{C}Wt}x(0))e^{\widehat{C}Wt}\text{diag}(x(0))W^\top dt
\end{aligned}$$

We again bound each of these two terms separately, as follows:

$$\begin{aligned}
&\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \left|\left|\int_0^\infty t\,\text{diag}((Q(t)+K^\top R(t)K)e^{\widehat{C}Wt}x(0))e^{\widehat{C}Wt}\text{diag}(x(0))W^\top dt\right|\right|_{\text{op}} \\
&\leq \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \int_0^\infty t\left|\left|\text{diag}((Q(t)+K^\top R(t)K)e^{\widehat{C}Wt}x(0))\right|\right|_{\text{op}}\left|\left|e^{\widehat{C}Wt}\right|\right|_{\text{op}}||\text{diag}(x(0))||_{\text{op}}\left|\left|W^\top\right|\right|_{\text{op}}dt \\
&= \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \int_0^\infty t\left|\left|(Q(t)+K^\top R(t)K)e^{\widehat{C}Wt}x(0)\right|\right|_{\infty}\left|\left|e^{\widehat{C}Wt}\right|\right|_{\text{op}}||x(0)||_\infty||W||_{\text{op}}dt \\
&\leq \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} \int_0^\infty t(\sqrt{n}\left|\left|Q(t)+K^\top R(t)K\right|\right|_\infty||W||_{\text{op}}||x(0)||_\infty^2)\left|\left|e^{\widehat{C}Wt}\right|\right|_{\text{op}}^2 dt.
\end{aligned}$$

Collecting all terms independent of t into a constant $D(K) = \max_t \sqrt{n}||Q(t)+K^\top R(t)K||_\infty||W||_{\text{op}}||x(0)||_\infty^2$ and using the bound $||e^{\widehat{C}Wt}|| \leq \beta(\widehat{C})e^{-\alpha(\widehat{C})t}$, we reach the conclusion as:

$$= \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))} D(K)\int_0^\infty t\left|\left|e^{\widehat{C}Wt}\right|\right|_{\text{op}}^2 dt \leq D(K)\beta(\widehat{C})^2\int_0^\infty te^{-2\alpha(\widehat{C})t}dt = \max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\frac{D(K)\beta(\widehat{C})^2}{4\alpha(\widehat{C})^2},$$

as desired. □

Once again, the proof in the finite time horizon case follows equivalently.

Corollary C.2. *[Deterministic, continuous-time, finite-horizon] Let* $J(K,C) := \int_0^T (x(t)^\top (Q(t) + K^\top R(t) K) x(t)) dt$ *for* $w = 0$. *Assume that* $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\Theta))) \geq 1 - \alpha$. *Then, under Assumption 3.1,*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\left(0 \leq \mathcal{R}(\Theta, C) \leq 2L\widehat{q} + \Delta_{\text{dom}}(\Theta, C)\right) \geq 1 - \alpha,$$

where L *is the Lipschitz constant of* $J(K, \widehat{C})$ *in* $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ *under the operator norm. Further, if* $\widehat{q} < r(C, K^*(C))$, *(see continuous-time in Definition A.1)* $\Delta_{\text{dom}}(\Theta, C) = 0$.

D Stochastic Discrete-Time Regret Analysis

Theorem D.1. *[Stochastic, discrete-time] Let* $J(K,C) := \sum_{t=0}^{\infty} (x_t^\top (Q_t + K^\top R_t K) x_t)$ *with* $w_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$ *across* t *such that* $D_2(K) := ||\Sigma||_{\text{op}} ||W||_{\text{op}} < \infty$. *Assume further that* $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\Theta))) \geq 1 - \alpha$. *Then, under Assumption 3.1 and Assumption 3.2,*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\left(0 \leq \mathcal{R}(\Theta, C) \leq 2L\widehat{q} + \Delta_{\text{dom}}(\Theta, C)\right) \geq 1 - \alpha,$$

where L *is the Lipschitz constant of* $J(K, \widehat{C})$ *in* $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ *under the operator norm. Further, if* $\widehat{q} < r(C, K^*(C))$, *(see discrete-time in Definition A.1)* $\Delta_{\text{dom}}(\Theta, C) = 0$.

Proof. We again consider any fixed θ and demonstrate the desired properties that $J(K, \widehat{C})$ is non-negative and L-Lipschitz in $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ under the operator norm for any $K \in \mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$, from which Lemma A.2 can be invoked to arrive at the desired conclusion. Notice the objective can be reformulated in the standard manner as follows:

$$\begin{aligned}
J(K, \widehat{C}) &:= \mathbb{E}[\sum_{t=0}^{\infty} (x_t^\top (Q_t + K^\top R_t K) x_t)] = \sum_{t=0}^{\infty} \mathbb{E}[(x_t^\top (Q_t + K^\top R_t K) x_t)] \\
&= \sum_{t=0}^{\infty} \mathbb{E}[\text{Tr}(x_t^\top (Q_t + K^\top R_t K) x_t)] = \sum_{t=0}^{\infty} \mathbb{E}[\text{Tr}((Q_t + K^\top R_t K) x_t x_t^\top)] \\
&= \sum_{t=0}^{\infty} \text{Tr}((Q_t + K^\top R_t K) \mathbb{E}[x_t x_t^\top]) = \sum_{t=0}^{\infty} \text{Tr}((Q_t + K^\top R_t K)(\mathbb{E}[x_t]\mathbb{E}[x_t]^\top + \text{Var}(x_t))).
\end{aligned}$$

We now use the following computations to evaluate this final expression:

$$\begin{aligned}
\mathbb{E}[x_t] &= \mathbb{E}[(\widehat{C}W) x_{t-1} + w_t] \\
&= \mathbb{E}[(\widehat{C}W) x_{t-1}] + \mathbb{E}[w_t] = \mathbb{E}[(\widehat{C}W)((\widehat{C}W) x_{t-2} + w_{t-1})] \\
&= \mathbb{E}[(\widehat{C}W)^2 x_{t-2}] + (\widehat{C}W)\mathbb{E}[w_{t-1}] = ... = (\widehat{C}W)^t x_0.
\end{aligned}$$

$$\begin{aligned}
\text{Var}(x_t) &= \text{Var}((\widehat{C}W) x_{t-1} + w_t) = (\widehat{C}W)\text{Var}(x_{t-1})(\widehat{C}W)^\top + \Sigma \\
&= (\widehat{C}W)\text{Var}(x_{t-2})(\widehat{C}W)^\top + (\widehat{C}W)\Sigma(\widehat{C}W)^\top + \Sigma = ... = \sum_{k=0}^{t-1} (\widehat{C}W)^k \Sigma (\widehat{C}W)^{k\top}.
\end{aligned}$$

With these simplifications, we are left with:

$$\begin{aligned}
J(K, \widehat{C}) = &\sum_{t=0}^{\infty} x_0^\top (\widehat{C}W)^{t\top} (Q_t + K^\top R_t K)(\widehat{C}W)^t x_0 \\
&+ \sum_{t=0}^{\infty} \sum_{k=0}^{t-1} \text{Tr}((Q_t + K^\top R_t K)(\widehat{C}W)^k \Sigma (\widehat{C}W)^{k\top})
\end{aligned}$$

J is clearly non-negative by construction. We demonstrate this quantity is Lipschitz continuous with an appropriate Lipschitz constant, again by bounding the gradient. The bound for the first term was demonstrated in the proof of Theorem B.1, for which reason we solely present that of the second term as follows:

$$\begin{aligned}
&\sum_{t=0}^{\infty}\sum_{k=0}^{t-1}\nabla_{\widehat{C}}\mathrm{Tr}((Q_t+K^\top R_tK)(\widehat{C}W)^k\Sigma(\widehat{C}W)^{k\top}))\\
&=\sum_{t=0}^{\infty}\sum_{k=0}^{t-1}k((((Q_t+K^\top R_tK)^\top(\widehat{C}W)^k\Sigma^\top)\odot(\widehat{C}W)^{(k-1)})W^\top+\\
&\sum_{t=0}^{\infty}\sum_{k=0}^{t-1}k(((Q_t+K^\top R_tK)(\widehat{C}W)^k\Sigma)\odot(\widehat{C}W)^{(k-1)})W^\top
\end{aligned}$$

We now bound each of these two terms separately, although the structure of the two is the same, so we explicitly show steps for bounding the first, from which the same can be repeated on the second.

$$\begin{aligned}
&\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}||\sum_{t=0}^{\infty}\sum_{k=0}^{t-1}k((((Q_t+K^\top R_tK)^\top(\widehat{C}W)^k\Sigma^\top)\odot(\widehat{C}W)^{(k-1)})W^\top||_{\mathrm{op}}\\
&\leq\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\sum_{t=0}^{\infty}\sum_{k=0}^{t-1}k||((Q_t+K^\top R_tK)^\top(\widehat{C}W)^k\Sigma^\top)\odot(\widehat{C}W)^{(k-1)}||_{\mathrm{op}}||W||_{\mathrm{op}}\\
&\leq\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\sum_{t=0}^{\infty}\sum_{k=0}^{t-1}k||Q_t+K^\top R_tK||_{\mathrm{op}}||(\widehat{C}W)^k||_{\mathrm{op}}||\Sigma||_{\mathrm{op}}||(\widehat{C}W)^{(k-1)}||_{\mathrm{op}}||W||_{\mathrm{op}}\\
&\leq\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\sum_{t=0}^{\infty}\sum_{k=0}^{t-1}k\beta_1e^{-\alpha_1t}||(\widehat{C}W)^k||_{\mathrm{op}}||\Sigma||_{\mathrm{op}}||(\widehat{C}W)^{(k-1)}||_{\mathrm{op}}||W||_{\mathrm{op}}
\end{aligned}$$

We again repeat the proof technique leveraged in Appendix B, where we first collect all terms independent of t into a constant $D_2(K)=||\Sigma||_{\mathrm{op}}||W||_{\mathrm{op}}$. By precisely the same argument as presented there, namely that K stabilizes $\widehat{C}$, we have that there is a $P\succ 0$ such that for some $\tau\in(0,1)$, we have

$$(\widehat{C}W)^\top P(\widehat{C}W)\preceq\tau^2P.\implies||(\widehat{C}W)^t||_{\mathrm{op}}\leq\kappa(P)\tau^t,$$

where $\kappa(P)$ is the condition number of P. By Assumption 3.2, we have that $\alpha_2(\widehat{C})\leq-\log(\tau)$ or $e^{-\alpha_2(\widehat{C})}\geq\tau$. Therefore, we have that the prior sum is bounded by

$$\begin{aligned}
\leq D_2(K)\kappa(P)^2\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\sum_{t=0}^{\infty}\beta_1e^{-\alpha_1t}\sum_{k=0}^{t-1}k\tau^{2k-1}&\leq D_2(K)\kappa(P)^2\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\sum_{t=0}^{\infty}\beta_1e^{-\alpha_1t}\sum_{k=0}^{t-1}ke^{-\alpha_2(\widehat{C})(2k-1)}\\
&\leq D_2(K)\kappa(P)^2\beta_1\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\sum_{k=0}^{\infty}ke^{-\alpha_2(\widehat{C})(2k-1)}\sum_{t=k}^{\infty}e^{-\alpha_1t}\\
&=D_2(K)\kappa(P)^2\beta_1\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\sum_{k=0}^{\infty}ke^{-\alpha_2(\widehat{C})(2k-1)}\left(\frac{e^{\alpha_1-\alpha_1k}}{e^{\alpha_1}-1}\right)\\
&=D_2(K)\kappa(P)^2\beta_1\left(\frac{1}{e^{\alpha_1}-1}\right)\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\sum_{k=0}^{\infty}k(e^{-\alpha_2(\widehat{C})})^{(2k-1)}(e^{-\alpha_1})^{k-1}\\
&=D_2(K)\kappa(P)^2\beta_1\left(\frac{1}{e^{\alpha_1}-1}\right)\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}e^{-\alpha_2(\widehat{C})}\sum_{k=0}^{\infty}k(e^{-(2\alpha_2(\widehat{C})+\alpha_1)})^{k-1}\\
&=D_2(K)\kappa(P)^2\beta_1\left(\frac{1}{e^{\alpha_1}-1}\right)\max_{\widehat{C}\in\mathcal{B}_{\widehat{q}}(f(\theta))}\frac{e^{-\alpha_2(\widehat{C})}}{(1-e^{-(2\alpha_2(\widehat{C})+\alpha_1)})^2},
\end{aligned}$$

thus completing the proof. □

E Stochastic Continuous-Time Regret Analysis

We introduce below the continuous-time analog of the discrete-time decay assumption made in Assumption 3.2 below.

Assumption E.1. For any θ, $\exists$ constants $\alpha_1, \beta_1 > 0$ such that for all $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$, $K \in \mathcal{K}(\widehat{C})$, and $t \geq 0$, $||Q(t) + K^\top R(t)K|| \leq \beta_1 e^{-\alpha_1 t}$ and $\min_{\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))}(2\alpha_2(\widehat{C}) + \alpha_1) > 0$ where $\alpha_2(\widehat{C}) := \inf_{K \in \mathcal{K}(\widehat{C})}(-\max_i \mathrm{Re}(\lambda_i(\widehat{C}W))) > 0$.

Theorem E.2. *[Stochastic, continuous-time] Let $J(K, C) := \mathbb{E}\left[\int_0^\infty (x(t)^\top (Q(t) + K^\top R(t)K)x(t))dt\right]$ with $w(t)$ a white noise process with spectral density Σ such that $D_2(K) := ||\Sigma||_{\mathrm{op}}||W||_{\mathrm{op}} < \infty$. Assume further that $\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{B}_{\widehat{q}}(f(\Theta))) \geq 1 - \alpha$. Then, under Assumption 3.1 and Assumption E.1,*

$$\mathcal{P}_{\Theta,C,\widehat{q}}\left(0 \leq \mathcal{R}(\Theta, C) \leq 2L\widehat{q} + \Delta_{\mathrm{dom}}(\Theta, C)\right) \geq 1 - \alpha,$$

where L is the Lipschitz constant of $J(K, \widehat{C})$ in $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ under the operator norm. Further, if $\widehat{q} < r(C, K^*(C))$, (see continuous-time in Definition A.1) $\Delta_{\mathrm{dom}}(\Theta, C) = 0$.

Proof. We again consider any fixed θ and demonstrate the desired properties that $J(K, \widehat{C})$ is non-negative and L-Lipschitz in $\widehat{C} \in \mathcal{B}_{\widehat{q}}(f(\theta))$ under the operator norm for any $K \in \mathcal{K}(\mathcal{B}_{\widehat{q}}(f(\theta)))$, from which Lemma A.2 can be invoked to arrive at the desired conclusion. Notice the objective can be reformulated in the standard manner as follows:

$$\begin{aligned}
J(K, \widehat{C}) &:= \mathbb{E}\left[\int_0^\infty (x(t)^\top (Q(t) + K^\top R(t)K)x(t))dt\right] \\
&= \int_0^\infty \mathbb{E}[(x(t)^\top (Q(t) + K^\top R(t)K)x(t))]dt = \int_0^\infty \mathbb{E}[\mathrm{Tr}(x(t)^\top (Q(t) + K^\top R(t)K)x(t))]dt \\
&= \int_0^\infty \mathbb{E}[\mathrm{Tr}((Q(t) + K^\top R(t)K)x(t)x(t)^\top)]dt = \int_0^\infty \mathrm{Tr}((Q(t) + K^\top R(t)K)\mathbb{E}[x(t)x(t)^\top])dt \\
&= \int_0^\infty \mathrm{Tr}((Q(t) + K^\top R(t)K)(\mathbb{E}[x(t)]\mathbb{E}[x(t)]^\top + \mathrm{Var}(x(t))))dt.
\end{aligned}$$

We now obtain the expressions for $\mathbb{E}[x(t)]$ and $\mathrm{Var}(x(t))$ using standard results from stochastic differential equations. For a full review on this topic, see (Särkkä and Solin, 2019):

$$\begin{aligned}
J(K, \widehat{C}) = &\int_0^\infty x(0)^\top (e^{\widehat{C}Wt})^\top (Q(t) + K^\top R(t)K)(e^{\widehat{C}Wt})x(0)dt \\
&+ \int_0^\infty \int_0^t \mathrm{Tr}((Q(t) + K^\top R(t)K)e^{\widehat{C}Wk}\Sigma e^{\widehat{C}Wk^\top})dkdt
\end{aligned}$$

J is clearly non-negative by construction. We demonstrate this quantity is Lipschitz continuous with an appropriate Lipschitz constant, again by bounding the gradient in much the same manner as the above bound. The bound for the first term was demonstrated in the proof of Theorem B.1, which holds under the finiteness assumption of $D(K)$. We, thus, solely present that of the second term as follows:

$$\begin{aligned}
&\int_0^\infty \int_0^t \nabla_{\widehat{C}} \mathrm{Tr}((Q(t) + K^\top R(t)K)e^{(\widehat{C}W)k}\Sigma e^{(\widehat{C}W)k^\top})dkdt \\
&= \int_0^\infty \int_0^t k(((Q(t)^\top + (R(t)K)^\top K)e^{k\widehat{C}W}\Sigma^\top) \odot e^{k\widehat{C}W})W^\top dkdt \\
&+ \int_0^\infty \int_0^t k(((Q(t) + K^\top R(t)K)e^{k\widehat{C}W}\Sigma) \odot e^{k\widehat{C}W})W^\top
\end{aligned}$$

We now bound each of these two terms separately, although the structure of the two is the same, so we explicitly show steps for bounding the first, from which the same can be repeated on the second.

$$\begin{aligned}
L &\leq \max_{\widehat{C}} || \int_0^\infty \int_0^t k(((Q(t)^\top + (R(t)K)^\top K)e^{k\widehat{C}W}\Sigma^\top) \odot e^{k\widehat{C}W})W^\top dkdt||_{\mathrm{op}} \\
&\leq \max_{\widehat{C}} \int_0^\infty \int_0^t k||Q(t) + K^\top R(t)K||_{\mathrm{op}}||\Sigma||_{\mathrm{op}}||W||_{\mathrm{op}}||e^{k\widehat{C}W}||_{\mathrm{op}}^2 dkdt
\end{aligned}$$

We again now collect all terms independent of t into a constant $D_2(K) = ||\Sigma||_{\text{op}}||W||_{\text{op}}$, leaving

$$\begin{aligned}
&\leq \max_{\widehat{C}} D_2(K) \int_0^\infty \beta_1 e^{-\alpha_1 t} \int_0^t k\beta_2(\widehat{C})^2 e^{-2\alpha_2(\widehat{C})k} dk dt \\
&= \max_{\widehat{C}} \frac{D_2(K)\beta_1\beta_2(\widehat{C})^2}{4\alpha_2^2(\widehat{C})} \int_0^\infty e^{-\alpha_1 t} \left(1 - 2\alpha_2(\widehat{C})te^{-2\alpha_2(\widehat{C})t} - e^{-2\alpha_2(\widehat{C})t}\right) dt \\
&= \max_{\widehat{C}} \frac{D_2(K)\beta_1\beta_2(\widehat{C})^2}{4\alpha_2^2(\widehat{C})} \left(\frac{1}{\alpha_1} - \frac{2\alpha_2(\widehat{C})}{(\alpha_1 + 2\alpha_2(\widehat{C}))^2} - \frac{1}{\alpha_1 + 2\alpha_2(\widehat{C})}\right) = \max_{\widehat{C}} \frac{D_2(K)\beta_1\beta_2(\widehat{C})^2}{\alpha_1(\alpha_1 + 2\alpha_2(\widehat{C}))^2}
\end{aligned}$$

We, therefore, again have the desired upper bound on the Lipschitz constant, as desired. □

F Unimodal Assumption Explanation

In classical engineering design, one would prescribe the dynamics of the system by explicitly writing out the physics of the system; this is so universally done that it may not even feel like an assumption in engineering design. This is the sense in which we mean that the design parameters commonly have some "unimodal" (often Dirac) measure in mapping to the system dynamics. For instance, if one is studying a cart pole system with a position x and angle θ, a common characterization (see Chapter 3 of (Tedrake, 2023)) is given by:

$$\begin{aligned}
\ddot{x} =& \frac{1}{m_c + m_p \sin^2\theta} \left[f_x + m_p \sin\theta(l\dot{\theta}^2 + g\cos\theta)\right] \\
\ddot{\theta} =& \frac{1}{l(m_c + m_p \sin^2\theta)} \left[-f_x \cos\theta - m_p l\dot{\theta}^2 \cos\theta \sin\theta - (m_c + m_p)g\sin\theta\right]
\end{aligned}$$

Here, the parameters m_c, m_p, and l could all be viewed as the "design parameters" θ depending on what one, as an engineer, has control over. Knowing that the underlying reality can be described by single set of dynamics is, therefore, what motivates using a unimodal model, as we wished to highlight with the models used in previous UCCD works.

G Coverage Guarantees Under Noisy Observations

Theorem G.1. *Let $\widetilde{C} = C + \epsilon$ where $vec(\epsilon) \sim \mathcal{N}(0, \Sigma)$, where $\epsilon \perp\!\!\!\perp (\Theta, C)$. Assume $\mathcal{U}(\theta) = \{C' \mid ||f(\theta) - C'||_{\text{op}} \leq \widehat{q}\}$ satisfies $\mathcal{P}_{\Theta,\widetilde{C},\widehat{q}}(\widetilde{C} \in \mathcal{U}(\Theta)) \geq 1 - \alpha$, where $||\cdot||_{\text{op}}$ denotes the matrix operator norm. If for any $\theta \in \Theta$ and $\delta > 0$, $\mathcal{P}(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta) > \mathcal{P}(\widehat{q}^2 \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta)$, then*

$$\mathcal{P}_{\Theta,C,\widehat{q}}(C \in \mathcal{U}(\Theta)) \geq \mathcal{P}_{\Theta,\widetilde{C},\widehat{q}}(\widetilde{C} \in \mathcal{U}(\Theta)) \geq 1 - \alpha.$$

Proof. Given that $\mathcal{P}_{\Theta,\widetilde{C},\widehat{q}}(\widetilde{C} \in \mathcal{U}(\Theta)) \geq 1-\alpha$, it suffices to show that $\mathcal{P}(C \in \mathcal{U}(\theta) \mid \Theta = \theta) \geq \mathcal{P}(\widetilde{C} \in \mathcal{U}(\theta) \mid \Theta = \theta)$ for all θ, as the conclusion can be drawn by the law of total probability:

$$\begin{aligned}
&\mathcal{P}(\widetilde{C} \in \mathcal{U}(\Theta) \mid \Theta = \theta) \\
&= \mathcal{P}\left(\left\|\widetilde{C} - f(\theta)\right\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) \\
&= \mathcal{P}\left(\sup_{\|x\|=1} \left\{\|Cx + \epsilon x - f(\theta)x\|_2^2\right\} \leq \widehat{q}^2 \mid \Theta = \theta\right) \\
&= \mathcal{P}\left(\sup_{\|x\|=1} \left\{\|Cx - f(\theta)x\|_2^2 + 2x^T\epsilon^T\left(Cx - f(\theta)x\right) + \|\epsilon x\|_2^2\right\} \leq \widehat{q}^2 \mid \Theta = \theta\right)
\end{aligned}$$

We now *lower* bound this inner quantity, from which

$$\begin{aligned}
&\sup_{\|x\|=1} \left\{ \|Cx - f(\theta)x\|_2^2 + 2x^T\epsilon^T \left(Cx - f(\theta)x\right) + \|\epsilon x\|_2^2 \right\} \\
&\quad \geq \sup_{\|x\|=1} \left\{ \|Cx - f(\theta)x\|_2^2 + 2x^T\epsilon^T \left(Cx - f(\theta)x\right) \right\} \\
&\quad \geq \|Cx' - f(\theta)x'\|_2^2 + 2(x')^T\epsilon^T \left(Cx' - f(\theta)(x')\right),
\end{aligned}$$

for any choice of $x' : ||x'||_2 = 1$. The second line follows by the trivial fact that $\|\epsilon x\|_2^2 \geq 0$ and the third from the fact that the previous line is the supremum of *all* such possible values x'. We now specifically take $x' := \arg\max_x \|Cx - f(\theta)x\|_2^2$ and denote it as x^*. From here, we arrive at the final bound

$$\begin{aligned}
&\sup_{\|x\|=1} \left\{ \|Cx - f(\theta)x\|_2^2 + 2x^T\epsilon^T \left(Cx - f(\theta)x\right) + \|\epsilon x\|_2^2 \right\} \\
&\quad \geq \|(C - f(\theta))(x^*)\|_2^2 + 2(x^*)^T\epsilon^T \left(Cx^* - f(\theta)(x^*)\right) \\
&\quad =: \|C - f(\theta)\|_{\text{op}}^2 + 2(x^*)^T\epsilon^T \left(Cx^* - f(\theta)(x^*)\right)
\end{aligned}$$

Since this is a *lower* bound on the original quantity of interest, we have that the probability this quantity is upper bounded by $\widehat{q}^2$ is *greater* than that of the original quantity being upper bounded. That is,

$$\begin{aligned}
&\mathcal{P}\left(\sup_{\|x\|=1} \left\{ \|Cx - f(\theta)x\|_2^2 + 2x^T\epsilon^T \left(Cx - f(\theta)x\right) + \|\epsilon x\|_2^2 \right\} \leq \widehat{q}^2 \mid \Theta = \theta \right) \\
&\leq \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 + 2{x^*}^T\epsilon^T \left(Cx^* - f(\theta)x^*\right) \leq \widehat{q}^2 \mid \Theta = \theta \right)
\end{aligned}$$

Let $\delta = 2{x^*}^T\epsilon^T \left(Cx^* - f(\theta)x^*\right)$. Then:

$$\begin{aligned}
&:= \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 + \delta \leq \widehat{q}^2 \mid \Theta = \theta \right) \\
&= \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 + \delta \leq \widehat{q}^2 \mid \Theta = \theta, \delta > 0 \right) \mathcal{P}(\delta > 0 \mid \Theta = \theta) \\
&\quad + \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 + \delta \leq \widehat{q}^2 \mid \Theta = \theta, \delta \leq 0 \right) \mathcal{P}(\delta \leq 0 \mid \Theta = \theta) \\
&= \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 - \delta \mid \Theta = \theta, \delta > 0 \right) \mathcal{P}(\delta > 0 \mid \Theta = \theta) \\
&\quad + \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta, \delta > 0 \right) \mathcal{P}(\delta \leq 0 \mid \Theta = \theta)
\end{aligned}$$

In this final line, we made use of the fact that

$$\mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 - \delta \mid \Theta = \theta, \delta \leq 0 \right) = \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta, \delta > 0 \right),$$

which follows since the distribution of δ is symmetric about 0 by the symmetry of the distribution of ϵ. In particular, $\delta = f(C, \epsilon)$; since $C \perp\!\!\!\perp \epsilon$, the joint distributions $\mathcal{P}(C, \epsilon)$ and $\mathcal{P}(C, -\epsilon)$ are identical. Thus, the distribution of $\delta' = f(C, -\epsilon)$ matches that of δ. Using this, we add terms that sum to 0 as follows:

$$\begin{aligned}
&= \mathcal{P}\Big(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta \Big) \\
&\quad - \mathcal{P}(\delta > 0 \mid \Theta = \theta)\, \mathcal{P}\Big(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta, \delta > 0 \Big) \\
&\quad - \mathcal{P}(\delta \leq 0 \mid \Theta = \theta)\, \mathcal{P}\Big(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta, \delta \leq 0 \Big) \Bigg\} = 0 \\
&\quad + \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 - \delta \mid \Theta = \theta, \delta > 0 \right) \mathcal{P}(\delta > 0 \mid \Theta = \theta) \\
&\quad + \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta, \delta > 0 \right) \mathcal{P}(\delta \leq 0 \mid \Theta = \theta),
\end{aligned}$$

where these newly added terms will be used for manipulation subsequently. From here, we re-express this expression with expectations, where we again use the symmetry in δ in the second term:

$$\begin{aligned}
&= \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) \\
&\quad - \mathcal{P}\left(\delta > 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta) \mid \delta > 0\right] \\
&\quad - \mathcal{P}\left(\delta \leq 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta) \mid \delta > 0\right] \\
&\quad + \mathcal{P}\left(\delta > 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 - \delta \mid \Theta = \theta\right) \mid \delta > 0\right] \\
&\quad + \mathcal{P}\left(\delta \leq 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta\right) \mid \delta > 0\right].
\end{aligned}$$

With this rewrite, we can group terms and conclude using the stated assumption on the "peaking" structure of the probability in the prediction region:

$$\begin{aligned}
&= \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) \\
&\quad - \mathcal{P}\left(\delta > 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) - \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 - \delta \mid \Theta = \theta\right) \mid \delta > 0\right] \\
&\quad + \mathcal{P}\left(\delta \leq 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta\right) - \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) \mid \delta > 0\right] \\
&= \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) \\
&\quad - \mathcal{P}\left(\delta > 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}\left(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) \mid \delta > 0\right] \\
&\quad + \mathcal{P}\left(\delta \leq 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}\left(\widehat{q}^2 \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta\right) \mid \delta > 0\right].
\end{aligned}$$

We, therefore, have that $\mathcal{P}(\widetilde{C} \in \mathcal{U}(\theta) \mid \Theta = \theta) \leq \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) + \Delta$, where

$$\begin{aligned}
\Delta := &\mathcal{P}\left(\delta \leq 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}\left(\widehat{q}^2 \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta\right) \mid \delta > 0\right] \\
&- \mathcal{P}\left(\delta > 0 \mid \Theta = \theta\right) \mathbb{E}\left[\mathcal{P}\left(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) \mid \delta > 0\right]
\end{aligned}$$

By the assumption, we know that for all $\delta > 0$:

$$\mathcal{P}\left(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 \mid \Theta = \theta\right) > \mathcal{P}\left(\widehat{q}^2 \leq \|C - f(\theta)\|_{\text{op}}^2 \leq \widehat{q}^2 + \delta \mid \Theta = \theta\right).$$

We also know that $\mathcal{P}(\delta \leq 0 \mid \Theta = \theta) \leq \mathcal{P}(\delta \geq 0 \mid \Theta = \theta)$ since

$$\begin{aligned}
\mathcal{P}(\delta \leq 0 \mid \Theta = \theta) &= \mathcal{P}\left(2{x^*}^T \epsilon^T \left(Cx^* - f(\theta)x^*\right) \leq 0 \mid \Theta = \theta\right) \\
&= \mathbb{E}_{C|\Theta=\theta}\left[\mathcal{P}_{\epsilon|C=c,\Theta=\theta}\left({x^*}^T \epsilon^T \left(c - f(\theta)\right) x^* \leq 0\right)\right] \\
&= \mathbb{E}_{C|\Theta=\theta}\left[0.5\right] \\
&= 0.5
\end{aligned}$$

Therefore, $\mathcal{P}\left(\left\|\widetilde{C} - f(\theta)\right\|_{\text{op}} \leq \widehat{q} \mid \Theta = \theta\right) - \mathcal{P}\left(\|C - f(\theta)\|_{\text{op}} \leq \widehat{q} \mid \Theta = \theta\right) \leq \Delta \leq 0$ □

H LQR C Gradient

We follow the presentation of (Fazel et al., 2018) to provide the derivation of $\nabla_C J(K, C)$. Note that the following derivation is given for the discrete-time setting; the continuous-time derivation follows in a similar fashion with a modification in the Lyapunov equations.

Lemma H.1. *Let $J(K, C)$ be the infinite horizon, discrete-time, deterministic analog of that defined in Equation (2), i.e. $J(K, C) := \sum_{t=0}^{\infty}(x_t^\top (Q_t + K^\top R_t K)x_t)$ for $w = 0$. Then,*

$$\nabla_C J(K, C) = 2P_K C W X_K W^T, \tag{11}$$

where X_K and P_K respectively solve the following two Lyapunov equations: $\Delta_K X_K \Delta_K^\top - X_K = -X_0$ and $P_K = \Delta_K^\top P_K \Delta_K + Q + K^\top R K$, where $\Delta_K := A - BK$.

Proof. By the standard reformulation of $J(K, C)$ as described in (Fazel et al., 2018), we can rewrite $J(K, C, x_0) = x_0^\top P_K x_0$, where we now make the notational change to make explicit the dependence on x_0, as it pertains to the derivation below. We then have that

$$\begin{aligned} J(K, C, x_0) &= x_0^\top \Delta_K^\top P_K \Delta_K x_0 + x_0^\top (Q + K^\top R K)x_0 \\ &= J(K, C, \Delta_K x_0) + x_0^\top (Q + K^\top R K)x_0. \end{aligned}$$

From here, we have that

$$\begin{aligned} \nabla_C J(K, C, x_0) &= \nabla_C J(K, C, \Delta_K x_0) + \nabla_C (x_0^\top (Q + K^\top R K)x_0) \quad (\to 0) \\ &= 2P_K C W x_0 x_0^T W^T + \nabla_C J(K, C, \Delta_K x_1)|_{x_1 := (A-BK)x_0} \\ &= ... \\ &= 2P_K C W (\sum_{t=0}^{\infty} x_t x_t^T) W^T \\ &= 2P_K C W X_K W^T, \end{aligned}$$

where the final equality follows from the well-known correspondence between this infinite sum and the aforementioned Lyapunov reformulation. □

I Policy Gradient Convergence Guarantee

Lemma I.1. *Suppose $f(x, y)$ is $c(y)$-gradient dominated for any $y \in \mathcal{Y}$, i.e. for any fixed y, there is a $c(y)$ such that, for any $x \in \mathcal{X}$ and $x^*(y) := \arg\min_{x \in \mathcal{X}} f(x, y)$:*

$$f(x, y) - f(x^*(y), y) \leq c(y)||\nabla_x f(x, y)||_F^2.$$

Further, let $\phi(x) := \max_{y \in \mathcal{Y}} f(x, y)$ and $x^* := \arg\min_{x \in \mathcal{X}} \phi(x)$. Assume that $\operatorname{Arg\,max}_{y \in \mathcal{Y}} f(x, y) \neq \emptyset \ \forall x$. Then,

$$\phi(x) - \phi(x^*) \leq c^*(x) \min_{y^* \in \operatorname{Arg\,max}_{y \in \mathcal{Y}} f(x,y)} ||\nabla_x f(x, y^*)||_F^2,$$

where $c^*(x) := \sup_{y^* \in \operatorname{Arg\,max}_{y \in \mathcal{Y}} f(x,y)} c(y^*)$.

Proof. Note that, for any maximizer $y^* \in \operatorname{Arg\,max}_{y \in \mathcal{Y}} f(x, y)$, we have

$$\begin{aligned} \phi(x) - \phi(x^*) &:= \max_{y \in \mathcal{Y}} f(x, y) - \max_{y \in \mathcal{Y}} f(x^*, y) = f(x, y^*) - \max_{y \in \mathcal{Y}} f(x^*, y) \\ &\leq f(x, y^*) - f(x^*, y^*) \leq f(x, y^*) - f(x^*(y^*), y^*) \leq c(y^*)||\nabla_x f(x, y^*)||_F^2 \end{aligned}$$

Given that this relationship holds for all such y^*, we immediately get that

$$\phi(x) - \phi(x^*) \leq \min_{y^* \in \operatorname{Arg\,max}_{y \in \mathcal{Y}} f(x,y)} c(y^*)||\nabla_x f(x, y^*)||_F^2 \leq c^*(x) \min_{y^* \in \operatorname{Arg\,max}_{y \in \mathcal{Y}} f(x,y)} ||\nabla_x f(x, y^*)||_F^2,$$

where $c^*(x) := \sup_{y^* \in \operatorname{Arg\,max}_{y \in \mathcal{Y}} f(x,y)} c(y^*)$. □

We now make use of the known fact that $J(K, C)$ is gradient-dominated for any fixed C, in turn satisfying the conditions of Lemma I.1, from which we reach the desired conclusion. The former fact was demonstrated in (Bu et al., 2019), which we present below for sake of convenience with modification of notational conventions to match that used herein.

Lemma I.2. *(Lemma 5 of (Fazel et al., 2018)) Let $J(K,C)$ be the infinite horizon, discrete-time, deterministic analog of that defined in Equation (2), i.e. $J(K,C) := \sum_{t=0}^{\infty}(x_t^\top(Q_t + K^\top R_t K)x_t)$ for $w = 0$. Then, if $X_K \succcurlyeq 0$ and $K \in \mathcal{K}(C)$,*

$$J(K,C) - J(K^*(C),C) \leq \frac{||X_{K^*(C)}||}{\sigma_{\min}(X_0)^2\sigma_{\min}(R)}||\nabla_K J(K,C)||_F^2, \tag{12}$$

where $X_0 \succ 0$ and X_K and P_K respectively solve the following two equations: $\Delta_K X_K \Delta_K^\top - X_K = -X_0$ and $P_K = \Delta_K^\top P_K \Delta_K + Q + K^\top RK$, where $\Delta_K := A - BK$.

Lemma I.3. *Let $\phi(K) := \max_{\widehat{C}\in\mathcal{C}} J(K,\widehat{C})$ and $K^*_{\text{rob}}(\mathcal{C}) := \arg\min_{K\in\mathcal{K}(\mathcal{C})}\phi(K)$, with J the infinite horizon, discrete-time, deterministic analog of that defined in Equation (2), i.e. $J(K,C) := \sum_{t=0}^{\infty}(x_t^\top(Q_t + K^\top R_t K)x_t)$ for $w = 0$. Then, for $K \in \mathcal{K}(\mathcal{C})$ where $X_K \succcurlyeq 0$ for all $\widehat{C} \in \mathcal{C}$, $\phi(K)$ satisfies*

$$\phi(K) - \phi(K^*_{\text{rob}}(\mathcal{C})) \leq \mu^*(K) \min_{C^*\in \text{Arg}\max_{\widehat{C}\in\mathcal{C}} J(K,\widehat{C})} ||\nabla_K J(K,C^*)||_F^2$$

for $\mu^*(K) := \sup_{C^*\in\text{Arg}\max_{\widehat{C}\in\mathcal{C}} J(K,\widehat{C})} \frac{||X_{K^*(C^*)}||}{\sigma_{\min}(X_0)^2\sigma_{\min}(R)}$, where $X_0 \succ 0$ and X_K and P_K respectively solve the following two equations: $\Delta_K X_K \Delta_K^\top - X_K = -X_0$ and $P_K = \Delta_K^\top P_K \Delta_K + Q + K^\top RK$, where $\Delta_K := A - BK$.

Proof. The proof for this follows immediately by demonstrating the assumption of Lemma I.1 is satisfied by Lemma I.2. □

Theorem I.4. *Let $\phi(K) := \max_{C\in\mathcal{C}} J(K,C)$ and $K^*_{\text{rob}}(\mathcal{C}) := \arg\min_{K\in\mathcal{K}(\mathcal{C})}\phi(K)$, with J the infinite horizon, discrete-time, deterministic analog of that defined in Equation (2), i.e. $J(K,C) := \sum_{t=0}^{\infty}(x_t^\top(Q_t + K^\top R_t K)x_t)$ for $w = 0$. Let $K^{(t)}$ be the t-th iterate of Algorithm 1. Assume for each iterate t, the optimization over C converges, i.e. $C^{(T_C)} = C^*(K^{(t)})$, that $K^{(t)} \in \mathcal{K}(\mathcal{C})$, and that $X_K \succcurlyeq 0$ for all $\widehat{C} \in \mathcal{C}$ and $K \in \mathcal{K}(\mathcal{C})$. Denote $\nu := \min_{\widehat{C}\in\mathcal{C}}\min_{K\in\mathcal{K}(\mathcal{C})}\sigma_{\min}(X_K)$. If in Algorithm 1*

$$\eta_K \leq \min_{[\widehat{A},\widehat{B}]:=\widehat{C}\in\mathcal{C}} \frac{1}{16}\min\Bigg\{\left(\frac{\sigma_{\min}(Q)\nu}{J(K,\widehat{C})}\right)^2 \frac{1}{||\widehat{B}||||\nabla_K J(K,\widehat{C})||(1+||\widehat{A}-\widehat{B}K||)}, \frac{\sigma_{\min}(Q)}{2J(K,\widehat{C})||R+\widehat{B}^\top P_K \widehat{B}||}\Bigg\}. \tag{13}$$

then, there exists a $\gamma > 0$ such that $\phi(K^{(T)}) - \phi(K^*_{\text{rob}}(\mathcal{C})) \leq (1-\gamma)^T(\phi(K_0) - \phi(K^*_{\text{rob}}(\mathcal{C})))$.

Proof. We follow the proof strategy developed in (Fazel et al., 2018), specifically in their presentation of Lemma 24, in which we leverage the above developed gradient dominance result, namely that in Lemma I.3. We first note that Algorithm 1 is equivalent to performing subgradient descent over $\phi(K)$ if we assume convergence of the inner maximization over C, that is if $C^{(T_C)} = C^*(K^{(t)})$ for some $C^* \in \text{Arg}\max_{\widehat{C}\in\mathcal{C}} J(K^{(t)},\widehat{C})$. It, therefore, suffices to characterize subgradient descent, where $K^{(t+1)} := K^{(t)} - \eta_K g_t$.

To complete this proof, it suffices to demonstrate $\phi(K^{(t)}) - \phi(K^{(t+1)}) \geq \gamma(K^{(t)})||g_t||_F^2$ for some $\gamma(K^{(t)}) > 0$, since this along with subgradient dominance can be used to establish the desired convergence guarantees by first demonstrating this intermediate result:

$$\begin{aligned}\phi(K^{(t+1)}) - \phi(K^*_{\text{rob}}(\mathcal{C})) &= (\phi(K^{(t+1)}) - \phi(K^{(t)})) + (\phi(K^{(t)}) - \phi(K^*_{\text{rob}}(\mathcal{C})))\\ &\leq -\gamma(K^{(t)})||g_t||_F^2 + (\phi(K^{(t)}) - \phi(K^*_{\text{rob}}(\mathcal{C})))\\ &\leq \left(1 - \gamma(K^{(t)})/\mu^*(K^{(t)})\right)(\phi(K^{(t)}) - \phi(K^*_{\text{rob}}(\mathcal{C}))).\end{aligned}$$

To then demonstrate the final convergence, we can simply apply this result inductively as follows:

$$\begin{aligned}\phi(K^{(T)}) - \phi(K^*_{\text{rob}}(\mathcal{C})) &\leq \left(1 - \gamma(K^{(T)})/\mu^*(K^{(T)})\right)(\phi(K^{(T-1)}) - \phi(K^*_{\text{rob}}(\mathcal{C})))\\ &\leq \left(1 - \gamma(K^{(T)})/\mu^*(K^{(T)})\right)\left(1 - \gamma(K^{(T-1)})/\mu^*(K^{(T-1)})\right)(\phi(K^{(T-2)}) - \phi(K^*_{\text{rob}}(\mathcal{C})))\\ &\leq \ldots \leq (\phi(K_0) - \phi(K^*_{\text{rob}}(\mathcal{C})))\prod_{t=1}^{T}\left(1 - \gamma(K^{(t)})/\mu^*(K^{(t)})\right) \leq (1-\gamma)^T(\phi(K_0) - \phi(K^*_{\text{rob}}(\mathcal{C}))),\end{aligned}$$

where we take $\gamma := \min_t \gamma(K^{(t)})/\mu^*(K^{(t)})$. We now prove $\phi(K^{(t)}) - \phi(K^{(t+1)}) \geq \gamma(K^{(t)})||g_t||_F^2$. To do so, we leverage the result of combining Lemmas 3 and 24 from (Fazel et al., 2018), by which it was demonstrated that for any fixed dynamics C, there is a $\beta(C) > 0$ such that $J(K, C) - J(K', C) \geq \beta(C)||\nabla_K J(K, C)||_F^2$ if $K, K' \in \mathcal{K}(C)$, $X_K \succcurlyeq 0$, and if η satisfies:

$$\eta \leq \frac{1}{16} \min\left((\frac{\sigma_{\min}(Q)\nu(C)}{J(K,C)})^2 \frac{1}{||B||||\nabla_K J(K,C)||(1+||A-BK||)}, \frac{\sigma_{\min}(Q)}{2J(K,C)||R+B^\top P_K B||} \right),$$

where $\nu(C) := \min_{K \in \mathcal{K}(C)} \sigma_{\min}(X_K)$. The stability assumption is satisfied in assuming all iterates $K^{(t)} \in \mathcal{K}(\mathcal{C})$, as $\mathcal{K}(\mathcal{C}) \subset \mathcal{K}(C)$. $X_K \succcurlyeq 0$ is similarly true under the assumption that this property holds for all optimization iterates. The assumption on the learning rate is guaranteed for any $C \in \mathcal{C}$ under the assumption of Equation (13). To leverage this result, we must, therefore, re-express the quantity of interest into an expression with fixed dynamics:

$$\begin{aligned}
\phi(K^{(t)}) - \phi(K^{(t+1)}) &:= J(K^{(t)}, C^*(K^{(t)})) - J(K^{(t+1)}, C^*(K^{(t+1)})) \\
&\geq J(K^{(t)}, C^*(K^{(t)})) - J(K^{(t+1)}, C^*(K^{(t)})) \\
&\geq \beta(C^*(K^{(t)}))||\nabla_K J(K^{(t)}, C^*(K^{(t)}))||_F^2 \\
&= \beta(C^*(K^{(t)}))||g_t||_F^2.
\end{aligned}$$

Thus, taking $\gamma(K^{(t)}) := \beta(C^*(K^{(t)}))$ satisfies the desired property and completes the proof. □

J Experimental Controls Setup

As discussed in Section 4, the standard approach to "robustness via multiplicative noise" is non-data-driven specification of the perturbations anticipated upon deployment. They all, however, share the same standard structure of Equation (7), with differences being in the specification of the collection $\{\delta_i\}_{i=1}^p, \{\gamma_i\}_{i=1}^q, \{A_i\}_{i=1}^p$, and $\{B_i\}_{i=1}^q$, where $p = q = 2$ is used across experiments. We consider two strategies for the specification of $(\{A_i\}, \{B_i\})$ and three for that of $(\{\delta_i\}, \{\gamma_i\})$. For the former:

- **Random**
  - $A_i[j,k] \sim \mathcal{N}(0,1)$
  - $B_i[j,k] \sim \mathcal{N}(0,1)$
- **Random Row-Col**
  - $A_i[j,:] = A_i[:,k] = 1$ for $j, k \sim \text{Unif}([n])$
  - $B_i[j,:] = B_i[:,k] = 1$ for $j \sim \text{Unif}([n]), k \sim \text{Unif}([m])$

For the latter, the general strategy is to find those $\{\delta_i\}_{i=1}^p, \{\gamma_i\}_{i=1}^q$ that result in unstable dynamics when paired with the corresponding $(\{A_i\}, \{B_i\})$ for some choice of controller, which varies across the strategies considered. This in turn defines a problem such that, within some radius of misspecified dynamics that retain stability, the controller still performs well. These methods proceed by initializing $\delta_i^{(0)} = \gamma_i^{(0)} = \mathbf{1}$ and iteratively multiplicatively increasing each by some pre-defined factor ρ such that $\delta_i^{(t)} = \rho\delta_i^{(t-1)}$ and similarly for $\gamma^{(t)}$ until $J(A, B, \{A_i\}, \{B_i\}, \{\delta_i^*\}, \{\gamma_i^*\}, K) = \infty$ in

$$\begin{aligned}
J(A, B, \{A_i\}, \{B_i\}, \{\delta_i\}, \{\gamma_i\}, K) &:= \int_0^\infty (x^\top Q x + (Kx)^\top R(Kx))dt \\
\text{s.t.} \quad \dot{x} &= \left((A + \sum_{i=1}^p \delta_i A_i) - (B + \sum_{i=1}^q \gamma_i B_i)K \right) x.
\end{aligned} \tag{14}$$

The problem specifications, therefore, vary in the K used as the stopping criterion of Equation (14) and whether $\{\delta_i^*\}, \{\gamma_i^*\}$ are modified in the final specification as follows:

- **Critical**: Consider $K^{(t)} := \arg\min_K J(A, B, \{A_i\}, \{B_i\}, \{\delta_i^{(t)}\}, \{\gamma_i^{(t)}\}, K)$ in each iterate; Take $\{\delta_i := \delta_i^*\}, \{\gamma_i := \gamma_i^*\}$

- **Open-Loop Mean-Square Stable (Weak)**: Consider $K := \mathbf{0}$; Take $\{\delta_i := \nu\delta_i^*\}, \{\gamma_i := \nu\gamma_i^*\}$ for some $\nu \in (0, 1)$
- **Open-Loop Mean-Square Unstable**: Consider $K := \mathbf{0}$; Take $\{\delta_i := \delta_i^*\}, \{\gamma_i := \gamma_i^*\}$

All prediction models $\widehat{f} : \Theta \to (A, B)$ were multi-layer perceptrons implemented in PyTorch (Paszke et al., 2019) with optimization done using Adam (Kingma and Ba, 2014) with a learning rate of 10^{-3} over 1,000 training steps. Training such models required roughly 10 minutes using an Nvidia RTX 2080 Ti GPU for each experimental setup. Running the robust control optimization algorithm took roughly one hour for 1,000 design trials.

K Experimental Dynamical Systems Setup

We consider the following dynamical systems in the experiments. Note that parameters were drawn from normal distributions centered on the nominally reported values from the respective papers these dynamics were considered from.

K.1 Aircraft Control

We consider the experimental setup studied in (Chrif and Kadda, 2014), in which optimal control is sought on the deflection angles of an aircraft. In particular, we assume the dynamics are given by the following:

$$A = \begin{bmatrix} \gamma_\beta & \gamma_p & \gamma_r & 1 \\ L_\beta & L_p & L_r & 0 \\ N_\beta & N_p & N_r & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} \gamma_{\delta_r} & \gamma_{\delta_a} \\ L_{\delta_r} & L_{\delta_a} \\ N_{\delta_r} & N_{\delta_a} \\ 0 & 0 \end{bmatrix}, \qquad \theta := [\gamma, L, N] \in \mathbb{R}^{15}$$

The parameter sampling distributions are given in Table 2.

Table 2: Sampling of parameters for aircraft control task.

Parameter Group	Symbols	Distribution	Hyperparameter Sampling
γ coefficients	$\gamma_\beta, \gamma_p, \gamma_r, \gamma_{\delta_r}, \gamma_{\delta_a}$	$\mathcal{N}(\mu_\gamma, \Sigma_\gamma)$	$\mu_\gamma \sim \mathcal{U}([0,1]^5)$, $\Sigma_\gamma = AA^\top$, $A \sim \mathcal{U}([0,1]^{5\times5})$
L coefficients	$L_\beta, L_p, L_r, L_{\delta_r}, L_{\delta_a}$	$\mathcal{N}(\mu_L, \Sigma_L)$	$\mu_L \sim \mathcal{U}([0,1]^5)$, $\Sigma_L = AA^\top$, $A \sim \mathcal{U}([0,1]^{5\times5})$
N coefficients	$N_\beta, N_p, N_r, N_{\delta_r}, N_{\delta_a}$	$\mathcal{N}(\mu_N, \Sigma_N)$	$\mu_N \sim \mathcal{U}([0,1]^5)$, $\Sigma_N = AA^\top$, $A \sim \mathcal{U}([0,1]^{5\times5})$

K.2 Load Positioning Control

We consider the load-positioning system of (Ahmadi, Rahmani, and Shahmansoorian, 2023; Jiang et al., 2016). In this case, the dynamics are given by:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -\frac{d_L}{m_L} - \frac{d_L}{m_B} & \frac{k_B}{m_B} & \frac{d_B}{m_B} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{d_L}{m_B} & -\frac{k_B}{m_B} & -\frac{d_B}{m_B} \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ \frac{1}{m_L} + \frac{1}{m_B} \\ 0 \\ -\frac{1}{m_B} \end{bmatrix}, \qquad \theta := [m_B, m_L, d_L, k_B, d_B] \in \mathbb{R}^5$$

The parameter sampling distributions are given in Table 3.

Table 3: Sampling of parameters for load positioning task.

Parameter	Symbol	Distribution	Hyperparameter Sampling
Mass of body	m_B	$m_B = 1/u$	$u \sim \mathcal{U}(0.04,\ 0.0667)$
Mass of load	m_L	$m_L = 1/u$	$u \sim \mathcal{U}(0.3333,\ 1.0)$
Stiffness of body	k_B	$k_B = u \cdot m_B$	$u \sim \mathcal{U}(0.4,\ 1.3333)$
Damping of body	d_B	$d_B = u \cdot m_B$	$u \sim \mathcal{U}(0.004,\ 0.0667)$

K.3 Furuta Pendulum

We also consider the Furuta pendulum dynamical system given in (Arulmozhi and Victorie, 2022), in which the system dynamics were specified by

$$A = \frac{1}{J_T}\begin{bmatrix} 0 & 0 & J_T & 0 \\ 0 & 0 & 0 & J_T \\ 0 & \frac{1}{4}M_p^2 L_p^2 L_r^2 g & -(J_p + \frac{1}{4}m_p L_p^2)D_r & \frac{1}{2}m_p L_p L_r D_p \\ 0 & -\frac{1}{2}m_p L_p g(J_r + m_p L_r^2) & \frac{1}{2}m_p L_p L_r D_r & -(J_r + m_p L_r^2)D_p \end{bmatrix} \qquad B = \frac{1}{J_T}\begin{bmatrix} 0 \\ 0 \\ J_p + \frac{1}{4}m_p L_p^2 \\ -\frac{1}{2}m_p L_p L_r \end{bmatrix}$$

$$\theta := [M_p, m_p, L_p, L_r, J_T, J_p, J_r, D_p, D_r] \in \mathbb{R}^9$$

The parameter sampling distributions are given in Table 4.

Table 4: Sampling of parameters for Furuta pendulum task.

Parameter	**Symbol**	**Distribution**	**Hyperparameter Values**
Pendulum mass	M_p	$\lvert\mathcal{N}(\mu_{M_p}, \sigma^2_{M_p})\rvert$	$\mu_{M_p} = 0.024,\ \sigma_{M_p} \sim \mathcal{U}(0,1)$
Rotor mass	m_p	$\lvert\mathcal{N}(\mu_{m_p}, \sigma^2_{m_p})\rvert$	$\mu_{m_p} = 0.095,\ \sigma_{m_p} \sim \mathcal{U}(0,1)$
Pendulum length	L_p	$\lvert\mathcal{N}(\mu_{L_p}, \sigma^2_{L_p})\rvert$	$\mu_{L_p} = 0.129,\ \sigma_{L_p} \sim \mathcal{U}(0,1)$
Rotor length	L_r	$\lvert\mathcal{N}(\mu_{L_r}, \sigma^2_{L_r})\rvert$	$\mu_{L_r} = 0.085,\ \sigma_{L_r} \sim \mathcal{U}(0,1)$
Total inertia	J_T	$\lvert\mathcal{N}(\mu_{J_T}, \sigma^2_{J_T})\rvert$	$\mu_{J_T} = f(\mu_{m_p}, \mu_{L_r}, \mu_{J_r}, \mu_{J_p}),\ \sigma_{J_T} \sim \mathcal{U}(0,1)$
Pendulum inertia	J_p	$\lvert\mathcal{N}(\mu_{J_p}, \sigma^2_{J_p})\rvert$	$\mu_{J_p} = \frac{M_p L_p^2}{12},\ \sigma_{J_p} \sim \mathcal{U}(0,1)$
Rotor inertia	J_r	$\lvert\mathcal{N}(\mu_{J_r}, \sigma^2_{J_r})\rvert$	$\mu_{J_r} = \frac{m_p L_r^2}{12},\ \sigma_{J_r} \sim \mathcal{U}(0,1)$
Pendulum damping	D_p	$\lvert\mathcal{N}(\mu_{D_p}, \sigma^2_{D_p})\rvert$	$\mu_{D_p} = 0.0005,\ \sigma_{D_p} \sim \mathcal{U}(0,1)$
Rotor damping	D_r	$\lvert\mathcal{N}(\mu_{D_r}, \sigma^2_{D_r})\rvert$	$\mu_{D_r} = 0.0015,\ \sigma_{D_r} \sim \mathcal{U}(0,1)$

K.4 DC Microgrids

We additionally consider the LQR model of DC microgrids given in (Liu et al., 2023), in which the system dynamics were specified by

$$A = \begin{bmatrix} \frac{2(-u_0 d - NK_2 S)}{V_s d} & 0 & \frac{-2NK_4 S}{V_s d} & \frac{-4NK_5 S}{V_s d} & \frac{2u_0}{V_s} & 0 & 0 & 0 & 0 \\ 0 & \frac{2(-u_0 d - NK_3 S)}{V_s d} & \frac{4NK_4 S}{V_s d} & \frac{6NK_5 S}{V_s d} & 0 & \frac{2u_0}{V_s} & 0 & 0 & 0 \\ \frac{6NK_2 S}{V_s d} & \frac{4NK_3 S}{V_s d} & \frac{2(-u_0 d - NK_4 S)}{V_s d} & 0 & 0 & 0 & \frac{2u_0}{V_s} & 0 & 0 \\ \frac{-4NK_2 S}{V_s d} & \frac{-2NK_3 S}{V_s d} & 0 & \frac{2(-u_0 d - NK_5 S)}{V_s d} & 0 & 0 & 0 & \frac{2u_0}{V_s} & 0 \\ \frac{u_0}{V_t} & 0 & 0 & 0 & \frac{-u_0}{V_t} & 0 & 0 & 0 & 0 \\ 0 & \frac{u_0}{V_t} & 0 & 0 & 0 & \frac{-u_0}{V_t} & 0 & 0 & 0 \\ 0 & 0 & \frac{u_0}{V_t} & 0 & 0 & 0 & \frac{-u_0}{V_t} & 0 & 0 \\ 0 & 0 & 0 & \frac{u_0}{V_t} & 0 & 0 & 0 & \frac{-u_0}{V_t} & 0 \\ \frac{NRT}{FC_2^c} & \frac{-NRT}{FC_3^c} & \frac{NRT}{FC_4^c} & \frac{NRT}{FC_5^c} & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} \frac{C_2^t - C_2^c}{V_s/2} \\ \frac{C_3^t - C_3^c}{V_s/2} \\ \frac{C_4^t - C_4^c}{V_s/2} \\ \frac{C_5^t - C_5^c}{V_s/2} \\ \frac{C_2^c - C_2^t}{V_t} \\ \frac{C_3^c - C_3^t}{V_t} \\ \frac{C_4^c - C_4^t}{V_t} \\ \frac{C_5^c - C_5^t}{V_t} \\ 0 \end{bmatrix}$$

$$\theta := [V_s, V_t, S, d, N, K_2, K_3, K_4, K_5, C_2^c, C_3^c, C_4^c, C_5^c, C_2^t, C_3^t, C_4^t, C_5^t] \in \mathbb{R}^{17}$$

The parameter sampling distributions are given in Table 5.

Table 5: Sampling of parameters for DC microgrids task.

Parameter	Symbol	Distribution	Hyperparameter Values
Source voltage	V_s	$\mathcal{N}(\mu_{V_s}, \sigma^2_{V_s})$	$\mu_{V_s} = 40,\ \sigma_{V_s} = 26.67$
Terminal voltage	V_t	$\mathcal{N}(\mu_{V_t}, \sigma^2_{V_t})$	$\mu_{V_t} = 500,\ \sigma_{V_t} = 333.33$
Surface area	S	$\mathcal{N}(\mu_S, \sigma^2_S)$	$\mu_S = 24,\ \sigma_S = 16.00$
Diffusion coefficient	d	$\mathcal{N}(\mu_d, \sigma^2_d)$	$\mu_d = 1.27 \times 10^{-3},\ \sigma_d = 8.47 \times 10^{-4}$
Number of layers	N	$\mathcal{N}(\mu_N, \sigma^2_N)$	$\mu_N = 37,\ \sigma_N = 24.67$
Reaction rate constant 2	K_2	$\mathcal{N}(\mu_{K_2}, \sigma^2_{K_2})$	$\mu_{K_2} = 8.768 \times 10^{-10},\ \sigma_{K_2} = 5.845 \times 10^{-10}$
Reaction rate constant 3	K_3	$\mathcal{N}(\mu_{K_3}, \sigma^2_{K_3})$	$\mu_{K_3} = 3.222 \times 10^{-10},\ \sigma_{K_3} = 2.148 \times 10^{-10}$
Reaction rate constant 4	K_4	$\mathcal{N}(\mu_{K_4}, \sigma^2_{K_4})$	$\mu_{K_4} = 6.825 \times 10^{-10},\ \sigma_{K_4} = 4.550 \times 10^{-10}$
Reaction rate constant 5	K_5	$\mathcal{N}(\mu_{K_5}, \sigma^2_{K_5})$	$\mu_{K_5} = 5.897 \times 10^{-10},\ \sigma_{K_5} = 3.931 \times 10^{-10}$
Capacitance cell 2 (cathode)	C_2^c	$\mathcal{N}(\mu_{C_2^c}, \sigma^2_{C_2^c})$	$\mu_{C_2^c} = 1.0,\ \sigma_{C_2^c} = 0.667$
Capacitance cell 3 (cathode)	C_3^c	$\mathcal{N}(\mu_{C_3^c}, \sigma^2_{C_3^c})$	$\mu_{C_3^c} = 1.0,\ \sigma_{C_3^c} = 0.667$
Capacitance cell 4 (cathode)	C_4^c	$\mathcal{N}(\mu_{C_4^c}, \sigma^2_{C_4^c})$	$\mu_{C_4^c} = 1.0,\ \sigma_{C_4^c} = 0.667$
Capacitance cell 5 (cathode)	C_5^c	$\mathcal{N}(\mu_{C_5^c}, \sigma^2_{C_5^c})$	$\mu_{C_5^c} = 1.0,\ \sigma_{C_5^c} = 0.667$
Capacitance cell 2 (total)	C_2^t	$\mathcal{N}(\mu_{C_2^t}, \sigma^2_{C_2^t})$	$\mu_{C_2^t} = 1.0,\ \sigma_{C_2^t} = 0.667$
Capacitance cell 3 (total)	C_3^t	$\mathcal{N}(\mu_{C_3^t}, \sigma^2_{C_3^t})$	$\mu_{C_3^t} = 1.0,\ \sigma_{C_3^t} = 0.667$
Capacitance cell 4 (total)	C_4^t	$\mathcal{N}(\mu_{C_4^t}, \sigma^2_{C_4^t})$	$\mu_{C_4^t} = 1.0,\ \sigma_{C_4^t} = 0.667$
Capacitance cell 5 (total)	C_5^t	$\mathcal{N}(\mu_{C_5^t}, \sigma^2_{C_5^t})$	$\mu_{C_5^t} = 1.0,\ \sigma_{C_5^t} = 0.667$

K.5 Fusion Plant

We finally consider the terminal sliding-mode control of a fusion plant from (Kirgni and Wang, 2023), given by:

$$A = \begin{bmatrix} -\frac{\beta}{\Lambda} & \frac{\beta_1}{\Lambda} & \frac{\beta_2}{\Lambda} & \frac{\beta_3}{\Lambda} & \frac{\alpha_f\theta}{\Lambda} & \frac{\alpha_c\theta}{2\Lambda} & -\frac{\sigma_X\theta}{\nu\Sigma_f\Lambda} & 0 \\ \lambda_1 & -\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda_2 & 0 & -\lambda_2 & 0 & 0 & 0 & 0 & 0 \\ \lambda_3 & 0 & 0 & -\lambda_3 & 0 & 0 & 0 & 0 \\ \frac{\epsilon_f P_0}{\mu_f} & 0 & 0 & 0 & -\frac{\Omega}{\mu_f} & \frac{\Omega}{\mu_f} & 0 & 0 \\ \frac{(1-\epsilon_f)P_0}{\mu_c} & 0 & 0 & 0 & \frac{\Omega}{\mu_c} & \frac{2M+\Omega}{2\mu_c} & 0 & 0 \\ (\gamma_X\Sigma_f - \sigma_X X_0)\phi_0 P_0 & 0 & 0 & 0 & 0 & 0 & -(\lambda_X + \phi_0 P_0 \theta) & \lambda_I \\ \gamma_I \Sigma_f \phi_0 P_0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_I \end{bmatrix} \quad B = \begin{bmatrix} -\frac{\theta}{\Lambda} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\theta := [\alpha_c, \alpha_f, \beta, \beta_1, \beta_2, \beta_3, \Lambda, \lambda_I, \lambda_X, \lambda_1, \lambda_2, \lambda_3, \mu_f, \mu_c, \gamma_X, \gamma_I, \sigma_X, \Sigma_f, \nu, \epsilon_f, \Omega, M, \theta, P_0, \phi_0, X_0] \in \mathbb{R}^{26}$$

The parameter sampling distributions are given in Table 6.

Table 6: Sampling of parameters for fusion plant task. Note that some of the parameters, specifically θ, μ_c, ν, Ω, M, ϕ_0, X_0, were not ascribed values in the paper from which the dynamics were provided. These were assumed to be in normalized units for the simulation.

Parameter	Symbol	Distribution	Hyperparameter Values
Coolant reactivity coefficient	α_c	$\mathcal{N}(\mu_{\alpha_c}, \sigma^2_{\alpha_c})$	$\mu = -2.0,\ \sigma = 2.0$
Fuel reactivity coefficient	α_f	$\mathcal{N}(\mu_{\alpha_f}, \sigma^2_{\alpha_f})$	$\mu = -14.0,\ \sigma = 14.0$
Total delayed neutron fraction	β	$\mathcal{N}(\mu_{\beta}, \sigma^2_{\beta})$	$\mu = 0.0065,\ \sigma = 0.0065$
Delayed neutron precursor (group 1)	β_1	$\mathcal{N}(\mu_{\beta_1}, \sigma^2_{\beta_1})$	$\mu = 0.00021,\ \sigma = 0.00021$
Delayed neutron precursor (group 2)	β_2	$\mathcal{N}(\mu_{\beta_2}, \sigma^2_{\beta_2})$	$\mu = 0.00225,\ \sigma = 0.00225$
Delayed neutron precursor (group 3)	β_3	$\mathcal{N}(\mu_{\beta_3}, \sigma^2_{\beta_3})$	$\mu = 0.00404,\ \sigma = 0.00404$
Prompt neutron lifetime	Λ	$\mathcal{N}(\mu_{\Lambda}, \sigma^2_{\Lambda})$	$\mu = 2.1,\ \sigma = 2.1$
Iodine decay constant	λ_I	$\mathcal{N}(\mu_{\lambda_I}, \sigma^2_{\lambda_I})$	$\mu = 10.0,\ \sigma = 10.0$
Xenon decay constant	λ_X	$\mathcal{N}(\mu_{\lambda_X}, \sigma^2_{\lambda_X})$	$\mu = 2.9,\ \sigma = 2.9$
Decay const. neutron prec. group 1	λ_1	$\mathcal{N}(\mu_{\lambda_1}, \sigma^2_{\lambda_1})$	$\mu = 0.0124,\ \sigma = 0.0124$
Decay const. neutron prec. group 2	λ_2	$\mathcal{N}(\mu_{\lambda_2}, \sigma^2_{\lambda_2})$	$\mu = 0.0369,\ \sigma = 0.0369$
Decay const. neutron prec. group 3	λ_3	$\mathcal{N}(\mu_{\lambda_3}, \sigma^2_{\lambda_3})$	$\mu = 0.632,\ \sigma = 0.632$
Fuel heat capacity	μ_f	$\mathcal{N}(\mu_{\mu_f}, \sigma^2_{\mu_f})$	$\mu = 0.0263,\ \sigma = 0.0263$
Coolant heat capacity	μ_c	$\mathcal{N}(\mu_{\mu_c}, \sigma^2_{\mu_c})$	$\mu = 1.0,\ \sigma = 1.0$
Fission yield (xenon)	γ_X	$\mathcal{N}(\mu_{\gamma_X}, \sigma^2_{\gamma_X})$	$\mu = 0.003,\ \sigma = 0.003$
Fission yield (iodine)	γ_I	$\mathcal{N}(\mu_{\gamma_I}, \sigma^2_{\gamma_I})$	$\mu = 0.059,\ \sigma = 0.059$
Xenon absorption cross-section	σ_X	$\mathcal{N}(\mu_{\sigma_X}, \sigma^2_{\sigma_X})$	$\mu = 3.5 \times 10^{-18},\ \sigma = 3.5 \times 10^{-18}$
Fission cross-section	Σ_f	$\mathcal{N}(\mu_{\Sigma_f}, \sigma^2_{\Sigma_f})$	$\mu = 0.3358,\ \sigma = 0.3358$
Neutrons per fission	ν	$\mathcal{N}(\mu_{\nu}, \sigma^2_{\nu})$	$\mu = 1.0,\ \sigma = 1.0$
Power deposition fraction in fuel	ϵ_f	$\mathcal{N}(\mu_{\epsilon_f}, \sigma^2_{\epsilon_f})$	$\mu = 0.92,\ \sigma = 0.92$
Heat transfer coefficient	Ω	$\mathcal{N}(\mu_{\Omega}, \sigma^2_{\Omega})$	$\mu = 1.0,\ \sigma = 1.0$
Coolant mass	M	$\mathcal{N}(\mu_{M}, \sigma^2_{M})$	$\mu = 1.0,\ \sigma = 1.0$
Control reactivity	θ	$\mathcal{N}(\mu_{\theta}, \sigma^2_{\theta})$	$\mu = 1.0,\ \sigma = 1.0$
Nominal reactor power	P_0	$\mathcal{N}(\mu_{P_0}, \sigma^2_{P_0})$	$\mu = 3.0,\ \sigma = \sqrt{3.0}$
Neutron flux	ϕ_0	$\mathcal{N}(\mu_{\phi_0}, \sigma^2_{\phi_0})$	$\mu = 1.0,\ \sigma = 1.0$
Nominal xenon conc.	X_0	$\mathcal{N}(\mu_{X_0}, \sigma^2_{X_0})$	$\mu = 1.0,\ \sigma = 1.0$

L Additional Experimental Results

L.1 Raw Results

We here provide the raw regrets from Section 5.1 in Table 7 and the proportion of unstable dynamics in Table 8.

Table 7: Each of the results below are median normalized regrets over 1,000 i.i.d. test samples with median absolute deviations in parentheses. For clarity, we have not reported any cases with $> 80\%$ unstable cases (see Table 8 for respective percentages).

	Airfoil	Load Positioning	Furuta Pendulum	DC Microgrids	Fusion Plant
Random Critical	—	—	—	—	—
Random OL MSS (Weak)	0.091 (0.045)	—	—	—	—
Random OL MSUS	—	—	—	—	—
Row-Col Critical	—	—	—	—	—
Row-Col OL MSS (Weak)	0.101 (0.063)	—	—	—	—
Row-Col OL MSUS	0.104 (0.066)	—	—	—	—
CPC	**0.085 (0.058)**	**0.033 (0.023)**	**0.002 (0.002)**	**0.000 (0.000)**	**0.011 (0.011)**
Shared Lyapunov	0.349 (0.221)	0.358 (0.255)	0.055 (0.039)	0.000 (0.000)	0.030 (0.027)
Auxiliary Stabilizer	0.322 (0.202)	0.343 (0.256)	0.048 (0.036)	0.000 (0.000)	0.029 (0.027)
$\mathcal{H}_\infty$	0.288 (0.188)	0.087 (0.060)	0.063 (0.045)	0.012 (0.010)	0.035 (0.032)

Table 8: Percentages of cases with unstable robust control over 1,000 i.i.d. test samples.

	Airfoil	Load Positioning	Furuta Pendulum	DC Microgrids	Fusion Plant
Random Critical	1.000	1.000	1.000	1.000	1.000
Random OL MSS (Weak)	0.783	1.000	0.920	1.000	0.987
Random OL MSUS	0.825	1.000	0.961	1.000	0.990
Row-Col Critical	0.998	1.000	1.000	1.000	1.000
Row-Col OL MSS (Weak)	0.200	1.000	0.948	1.000	0.960
Row-Col OL MSUS	0.210	1.000	0.951	1.000	0.963
CPC	0.088	0.251	0.174	0.009	0.643
Shared Lyapunov	0.093	0.229	0.141	0.008	0.561
Auxiliary Stabilizer	0.087	0.223	0.142	0.007	0.556
$\mathcal{H}_\infty$	0.081	0.236	0.142	0.007	0.570

L.2 Method Timings

CPC is more computationally expensive than alternatives. This pairs well with the anticipated use cases, namely in engineering design workflows involving UCCD, i.e. where the control problem is solved *offline*.

Table 9: Comparison of average method timing (to convergence) across tasks as measured over 10 trials for each experimental setup.

Method	**Airfoil**	**Load Position**	**Pendulum**	**DC Microgrids**	**Fusion**
$\mathcal{H}_\infty$	0.17	0.14	0.16	0.19	0.23
Shared Lyapunov	2.68	2.63	2.53	1.13	2.12
Auxiliary Stabilizer	1.70	1.75	1.85	1.01	1.41
Random Critical	7.50	1.74	6.78	7.02	10.50
Random OL MSS (Weak)	6.72	1.34	4.94	7.36	6.39
Random OL MSUS	6.63	2.18	7.25	15.21	7.11
Row-Col Critical	5.37	2.32	5.51	5.48	5.44
Row-Col OL MSS (Weak)	4.35	2.34	4.13	1.00	2.88
Row-Col OL MSUS	3.43	1.84	4.77	4.54	3.18
CPC	13.03	13.44	12.47	12.19	10.88