CONFORMAL ROBUST CONTROL OF LINEAR SYSTEMS

Yash Patel
Department of Statistics
University of Michigan
Ann Arbor, MI 48104
yppatel@umich.edu

Sahana Rayan
Department of Statistics
University of Michigan
Ann Arbor, MI 48104
srayan@umich.edu

Ambuj Tewari
Department of Statistics
University of Michigan
Ann Arbor, MI 48104
tewaria@umich.edu

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 then demonstrate that such a controller can be efficiently produced via a novel policy gradient method that has convergence guarantees. We finally demonstrate the superior empirical performance of our method over alternate robust control specifications, such as $\mathcal{H}_\infty$ and LQR with multiplicative noise, across a collection of engineering control systems.

1 Introduction

Seeking control over a family of dynamical systems is a problem often encountered in engineering [1, 2, 3]. 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) [4]. Traditional engineering design loops operated sequentially, first proposing a design and then developing a controller [5, 6]. Such workflows, however, sacrificed the improved optimality possible in their coupling, hence the increasing interest in leveraging end-to-end co-control design pipelines [7, 8].

Initial works in CCD studied optimal design assuming perfectly specified, deterministic system dynamics [9, 10, 11, 12]. Such assumptions have, however, become overly restrictive, resulting in interest in robust extensions of the CCD formulation, referred to as uncertain CCD (UCCD) [13, 14, 15]. Such uncertainty can arise from many sources in the design process, such as noise in the control channels, uncertainties in the design parameters, or unmodeled components of the system dynamics. In addition, the nature of the UCCD problem specification differs depending on the risk tolerance in the downstream engineering application of interest. For instance, in risk-neutral settings, stochastic problem specifications are often appropriate [16, 17, 18], whereas in more risk-averse settings probabilistic [19, 20, 21] or worst-case problem specifications [22, 23, 24] are preferred.

We focus herein on the worst-case robust UCCD formulation (WCR-UCCD), specifically on dynamics misspecification. WCR-UCCD requires the specification of a dynamics uncertainty region. Existing methods of specification, however, tend to be ad-hoc, relying on intuition. Such methods, 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 [16].

We, thus, focus herein on providing a principled 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

[25, 26, 27]. LQR systems are of broad interest both due to their analytic tractability and widespread applicability to practical engineering systems, especially due to their extensibility to nonlinear systems via Koopman operators [28, 29, 30]. 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 [31]. 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 over infinite or finite time horizons and demonstrating empirical improvements over alternative robust control schemes.
- Extending the line of 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, efficient policy gradient algorithm for obtaining the optimal robust controller with convergence guarantees proven via gradient dominance.

2 Background

2.1 Conformal Prediction

Conformal prediction is a principled, distribution-free approach of uncertainty quantification [32, 33]. "Split conformal," the most common variant of conformal prediction, is used as a wrapper around black-box predictors $\widehat{f} : \mathcal{X} \to \mathcal{Y}$ such that prediction *regions* $\mathcal{C}(x)$ are returned in place of the typical point predictions $\widehat{f}(x)$. Prediction regions $\mathcal{C}(x)$ are specifically sought to have coverage guarantees on the true $y := f(x)$. That is, for some prespecified α, we wish to have $\mathcal{P}_{X,Y}(Y \in \mathcal{C}(X)) \geq 1 - \alpha$.

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. The training dataset is used in the standard manner to fit $\widehat{f}$. After fitting $\widehat{f}$, 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 $\mathcal{C}(x)$ to be $\{y \mid s(x, y) \leq \widehat{q}\}$. Such $\mathcal{C}(x)$ satisfies the aforementioned coverage guarantees 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 [33]. A practical objective of conformal prediction, thus, is to define score functions that maximize this predictive efficiency.

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 directly leveraged, i.e. taking $w^*(x) := \min_w f(w, \widehat{c})$. Such an approach, however, is inappropriate in safety-critical settings, given that the predictor function $\widehat{g}$ will likely be misspecified and, thus, may result in suboptimal decisions under the true cost parameter, which we denote as c. For this reason, robust alternatives to the nominal formulation have become of interest [34, 35, 36]. We focus on the formulation posited in [31], which studied

$$w^*(x) := \min_{w} \max_{\widehat{c} \in \mathcal{U}(x)} \quad f(w, \widehat{c}) \quad \text{s.t.} \quad \mathcal{P}_{X,C}(C \in \mathcal{U}(X)) \geq 1 - \alpha, \tag{1}$$

where $\mathcal{U} : \mathcal{X} \to \mathcal{F}$ is a uncertainty region predictor, with $\mathcal{F}$ being the σ-field of $\mathcal{C}$. Works in this field typically focus on both theoretically characterizing and empirically studying the resulting 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 [31], $\mathcal{U}(x)$ was specifically constructed via conformal prediction to provide probabilistic guarantees; that is, by taking $\mathcal{U}(x) := \mathcal{C}(x)$ to be the prediction region produced by conformalizing the predictor $\widehat{g}$, they demonstrated $\mathcal{P}_{X,C}\left(0 \leq \Delta(X, C) \leq L \operatorname{diam}(\mathcal{U}(X))\right) \geq 1 - \alpha$.

2.3 Linear Quadratic Regulators & Control Co-Design

The field of control has a long history in engineering physics and robotics [37]. The most heavily studied specific control setup is the linear quadratic regulator (LQR), where dynamics of the underlying state x are assumed to have a linear form, consisting of both an autonomous and controlled component, namely $\dot{x} = Ax + Bu + w$, where u are the control inputs and $w \sim \mathcal{D}_w$ the noise. Optimal control is then posed as a constrained optimization problem, with the objective $J(u)$ consisting of a term related to the deviation from a target state and another related to the control input. Under linear dynamics and quadratic cost, it is well-known that the optimal controller has a linear feedback form, namely $u^*(x) = -K^*x$ where K^* is known as the "optimal gain matrix." For this reason, optimization is sought over gain matrices K rather than a more general function space of controllers:

$$K^*(A,B) := \underset{K \in \mathcal{K}(A,B)}{\arg\min}\ \mathbb{E}[J(K,A,B)] \quad \text{where } J(K,A,B) := \int_0^\infty (x^\top Q x + (Kx)^\top R(Kx))dt$$
$$\text{s.t.} \quad \mathcal{K}(A,B) := \{K : \mathrm{Re}(\lambda_i(A - BK)) < 0 \quad \forall i\} \quad \dot{x} = (A - BK)x + w$$

$\mathcal{K}(A,B)$ is known as the set of "stabilizing controllers" for the A, B system dynamics. Variations, where the integral is replaced by a discretized sum or where the cost is only considered to some finite T time horizon, are also of interest in some settings. Solving this problem is typically done either by solving the algebraic Ricatti equation [38] or via policy gradient techniques [39].

We are interested herein in the robust control problem subproblem in the context of uncertain co-control design. We briefly summarize the relevant pieces of UCCD herein; for a full survey, refer to [13]. Engineering designs can often be specified by parameters θ, which could specify, 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 described by θ_1 will fly differently from one given by θ_2. Worst-case robust UCCD with dynamics misspecification, thus, solves:

$$\min_{K,\theta} \max_{\widehat{A} \in \mathcal{A}(\theta), \widehat{B} \in \mathcal{B}(\theta)} \mathbb{E}[o(K, \widehat{A}, \widehat{B}, \theta)] \qquad \text{s.t.} \quad \dot{x} = \widehat{A}x + \widehat{B}u + w, \tag{2}$$

where $(\mathcal{A}(\theta), \mathcal{B}(\theta))$ are uncertainty sets of the dynamics for such a design and o is the objective function of interest. 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. $o(K, A, B, \theta) := \ell(\theta) + J(K, A(\theta), B(\theta))$ [40, 25]. One of the most commonly applied solution techniques of Equation (2) is via bilevel optimization, in which an outer optimization loop is performed over design parameters and an inner one over controllers for the current design iterate [41, 42]. For this reason, the specification of the robust control subproblem can be studied independently of the outer design optimization loop. Often, as promising designs emerge, they are field tested. Doing so often then informs a more refined dynamics model $q(X \mid \theta)$ that can be used for further in-silico optimization [43, 44, 45].

3 Conformalized Predict-Then-Control

We now introduce an approach to define and solve a robust LQR problem with probabilistic guarantees, discussing the problem formulation in Section 3.1 and the probabilistic guarantees on the resulting synthesized controllers in Section 3.3 and Section 3.4. We then provide an algorithm to extract this controller with convergence guarantees in Section 3.5.

3.1 Problem Formulation

For the presentation below, we use the following notational conventions. Assume $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, $A \in \mathbb{R}^{n \times n}$, 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 $(\theta^{(i)}, C^{(i)}) \sim \mathcal{P}(\Theta, C)$ and $K^{(i)} \sim \mathcal{P}(K)$, 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:

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

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 as specified in Section 2.3 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 Section 2.3 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, $(\widetilde{A}^{(i)}, \widetilde{B}^{(i)})$ [46]. 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 $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$, 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 a matrix-norm residual, namely: $s(\theta, C) = ||f(\theta) - C||_2$, where $||\cdot||_2$ is the matrix *operator* norm, i.e. $||A||_2 = \sigma_{\max}(A)$, 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 $||\cdot||_2$ metric. We note that, in place of this, another matrix norm could be leveraged with minimal modification.

3.3 Coverage Guarantee Consequences

We now characterize the regret induced through the introduction of 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 an expectation over the *true* C dynamics.

The regret characterizations provided below are stated in terms of a function $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$, where $\Omega(\mathcal{C})$ denotes the space of all subsets of the $\mathcal{C}$ space. Such a g quantifies the regret resulting from being robust to the uncertainty region $\mathcal{U}(\theta) \in \Omega(\mathcal{C})$ under a true dynamic $C \in \mathcal{C}$. Such a suboptimality function should have the characteristics that $g(\{C\}, C) = 0$, i.e. there should be no suboptimality if $\mathcal{U}(\theta)$ is the singleton of the true dynamic C, and that it is monotonically increasing in set size for any fixed C. We refer to any such g as a "pinned monotonic set function."

Definition 3.1. $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$ is a *pinned monotonic set function* over $\mathcal{C}$ if $\forall C \in \mathcal{C}$, $g(\{C\}, C) = 0$, $g(\Omega, C) < \infty$ $\forall \Omega \in \Omega(\mathcal{C})$ such that $\text{diam}(\Omega) < \infty$, and $\Omega_1 \subset \Omega_2 \implies g(\Omega_1, C) \leq g(\Omega_2, C)$.

We provide the regret statements for the continuous, infinite time horizon case for both deterministic and stochastic dynamics below; the explicit form of g is given in the proof in Appendix A. Both settings require a mild assumption that the problem parameters have bounded norms, formalized in Assumption 3.2; this will hold for any properly specified problem setup. Notably, however, the two settings differ in the assumptions required on the discounting of $Q(t)$ and $R(t)$; the case of stochastic dynamics requires $Q(t)$ and $R(t)$ be discounted over t, while the deterministic case is fully compatible with non-discounted rewards. This discounting is given in Assumption 3.3.

Assumption 3.2. For any θ, $K \in \mathcal{K}(\mathcal{U}(\theta))$, and $C \in \mathcal{U}(\theta)$, $D(K) := \sqrt{n}||Q(t) + K^\top R(t) K||_\infty ||x_0||_\infty^2 ||W||_2 < \infty$.

Assumption 3.3. For any θ, $K \in \mathcal{K}(\mathcal{U}(\theta))$, and $C \in \mathcal{U}(\theta)$, there exist constants $\alpha_1, \beta_1 > 0$ such that $||Q(t) + K^\top R(t) K|| \leq \beta_1 e^{-\alpha_1 t}$ and $\min_C(2\alpha_2(C) + \alpha_1) > 0$ where $\alpha_2(C) := \max_{K \in \mathcal{K}(C)}(-\max_i \text{Re}(\lambda_i(CW)))$.

We highlight that significant proof modifications were required over the comparable proof of [31], stemming from the complications in handling the universal stabilizing set in the regret computation. Our setting also has the proof complication not present in [31] that the domain of the *minimization* is also a function of $\mathcal{U}(\theta)$, again through the universal stabilizing set. The full collection of statements and proofs across discrete- and continuous-time dynamics and infinite and finite time horizons are provided in Appendices B to E.

Theorem 3.4. *Let $J(K, C)$ be the infinite horizon, continuous-time, deterministic analog of that defined in Section 2.3, i.e. $J(K, C) := \int_0^\infty (x(t)^\top (Q + K^\top RK)x(t))dt$ for $w = 0$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. Then under Assumption 3.2, there is a pinned monotonic set function g over $\mathcal{C}$ that satisfies: $\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1 - \alpha$.*

Theorem 3.5. *Let $J(K, C)$ be the infinite horizon, continuous-time, stochastic analog of that defined in Section 2.3, i.e. $J(K, C) := \mathbb{E}\left[\int_0^\infty (x(t)^\top (Q(t) + K^\top R(t)K)x(t))dt\right]$ with $w(t)$ a white noise process with spectral density Σ such that $||\Sigma||_2||W||_2 < \infty$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1-\alpha$. Then under Assumption 3.2 and Assumption 3.3, there is a pinned monotonic set function g over $\mathcal{C}$ that satisfies: $\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1 - \alpha$.*

These regret statements demonstrate that the robust controller produces an upper bound to the value of the nominal control problem across designs whose looseness increases monotonically with the size of the prediction regions. 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.

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 linear time-invariant system identification in Theorem 9.1 of [46].

The coverage guarantee result given in Theorem 3.6 is the multivariate extension of Theorem A.5 from [47] 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, and remains true of data-driven surrogates in UCCD, highlighted by the use of neural-based point predictor surrogates [48, 49]. $\mathcal{U}(\theta)$ 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 for problems of interest in Section 5. The full proof of this theorem is deferred to Appendix F.

Theorem 3.6. *Let $\widetilde{C} = C + \epsilon$ where vec$(\epsilon) \sim \mathcal{N}(0, \Sigma)$. Assume $\mathcal{U}(\theta) = \{C' \mid ||f(\theta) - C'||_2 \leq \widehat{q}\}$ satisfies $\mathcal{P}_{\Theta,\widetilde{C}}(\widetilde{C} \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. If for any θ, $\mathcal{P}(C \mid \theta)$ peaks in $\mathcal{U}(\theta)$, that is for any $\delta > 0$, $\mathcal{P}_{C|\Theta=\theta}(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2) > \mathcal{P}_{C|\Theta=\theta}(\widehat{q}^2 \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2 + \delta)$, then*

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

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, in which $x_{t+1} = (A - BK)x_t$. We also assume that $x(0)$ is known deterministically; extension to cases where $x(0) \sim \mathcal{N}(0, \Sigma)$ is straightforward. 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, 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

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

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 $\phi(K) := \max_{\widehat{C} \in \mathcal{U}(\theta)} J(K, \widehat{C})$. By Danskin's Theorem, $\nabla_K \phi(K) = \nabla_K J(K, C^*(K))$, where $C^*(K) := \arg\max_{\widehat{C} \in \mathcal{U}(\theta)} J(K, \widehat{C})$ as long as $C^*(K)$ is the unique maximizer. This uniqueness assumption is unlikely to be an issue in practice when the dynamics predictor is sufficiently accurate and $\mathcal{U}(\theta)$ consequently small, as we empirically verify in Section 5.1.

Thus, policy gradient in the robust formulation proceeds through the evaluation of Equation (3) with $[A^*, B^*] := C^*(K)$ and (X_K^*, P_K^*), the solutions to the Lyapunov equations when $\Delta_K^* := A^* - B^* K$. 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 efficiently computed with gradient ascent, where the Lyapunov expression $\nabla_C J(K, C) = 2P_K C W X_K W^\top$ is derived in Appendix G. The use of Danskin's Theorem for policy gradient 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 (discussed in Section 4) could be directly translated into GAREs and, therefore, did not require algorithmic innovation. The algorithm is summarized in Algorithm 1.

Algorithm 1 Conformalized Predict-Then-Control (CPC)

Inputs: Design θ, Predictor $f(\theta)$, Conformal quantile $\widehat{q}$, Step sizes η_K, η_C, Max steps T_K, T_C
$K^{(0)} \sim U(\mathcal{K}), \widehat{C} := f(\theta)$
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**
    $\Delta^{(t_C)} := A^{(t_C)} - B^{(t_C)} K^{(t_K)}$
    $X^{(t_C)} \leftarrow \text{Solve}(\ell_X(X, \Delta^{(t_C)}) = 0; X)$
    $P^{(t_C)} \leftarrow \text{Solve}(\ell_P(P, \Delta^{(t_C)}, K^{(t_K)}) = 0; P)$
    $C^{(t_C+1)} \leftarrow \Pi_{\mathcal{B}_{\widehat{q}}(\widehat{C})}(C^{(t_C)} + \eta_C(2P^{(t_C)} C^{(t_C)} W^{(t_K)} X^{(t_C)} (W^{(t_K)})^\top))$
  $\Delta^* := A^* - B^* K^{(t_K)}$ ▷ $C^* := C^{(T_C)}$
  $X^* \leftarrow \text{Solve}(\ell_X(X, \Delta^*) = 0; X)$
  $P^* \leftarrow \text{Solve}(\ell_P(P, \Delta^*, K^{(t_K)}) = 0; P)$
  $K^{(t_K+1)} \leftarrow K^{(t_K)} - \eta_K(2((R + (B^*)^\top P^* B^*) K^{(t_K)} - (B^*)^\top P^* A^*) X^*)$
Return $K^{(T_K)}$

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 nominal case despite the nonconvexity of the problem in K due to a property known "gradient dominance" [50]. We say 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 first demonstrate that gradient dominance is retained in our robust formulation in Lemma H.3 and that this then produces convergence guarantees for Algorithm 1 for any fixed θ. Due again to the novelty of our problem framing, we highlight that the proof strategy leverages Danskin's Theorem in a novel manner to demonstrate gradient dominance.

The full proof is deferred to Appendix H and parallels the nominal proof presented in [51]; 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. In line with [51], 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 the specialization of $\mathcal{C} := \mathcal{U}(\theta)$. Note that we also defer the presentation of the explicit poly-expression in Theorem 3.7 to Appendix H.

Theorem 3.7. *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 Section 2.3, i.e. $J(K, C) := \sum_{t=0}^{\infty} (x_t^\top (Q + K^\top R K) x_t)$ 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* $\eta_K \leq \min_{[\widehat{A},\widehat{B}]\in\mathcal{C}} \text{poly}(\frac{\nu\sigma_{\min}(Q)}{J(K^{(0)},C)}, \frac{1}{||\widehat{A}||}, \frac{1}{||\widehat{B}||}, \frac{1}{||R||}, \sigma_{\min}(R))$, *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. Such a projection operation would require solving the following robust optimization problem:

$$\Pi_{\mathcal{K}(\mathcal{U}(\theta))}(\widetilde{K}^{(t)}) := \arg\min_K ||K - \widetilde{K}^{(t)}||_2 \quad \text{s.t.} \max_{[\widehat{A},\widehat{B}]:=\widehat{C}\in\mathcal{U}(\theta)} \max_i \text{Re}(\lambda_i(\widehat{A} + \widehat{B}K)) < 0.$$

There, however, is no known efficient algorithm to solve this projection step. Despite being of theoretical concern, this instability issue fails to be relevant in systems of practical relevance, since the controller iterates remain well within the set of stabilization in such cases for sufficiently well-calibrated 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 $\mathcal{U}(\theta)$. Characterizing such approximations is of interest in future works.

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 [13, 52, 53], whereas the latter directly incorporates desired robustness into the optimization problem *prior* to deployment [54, 55]. Notably, such robustness schemes are orthogonal: one can incorporate desired robustness in the initial control design and adapt the fitted controller upon deployment. Given that control co-design seeks to identify a controller *prior* to deploying a design, we, therefore, specifically highlight methods of trajectory-free robust control.

The most widely-known classical robust 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 analogously defining $z = \begin{bmatrix} Q^{1/2}x & R^{1/2}u \end{bmatrix}$, we recover a recognizable robust LQR formulation with a form identical to that in Section 2.3, with the objective replaced by

$$u^* = \min_{\{u_t\}} \max_{\{w_t\}} \sum_{t=0}^{\infty} (x_t^\top Q x_t + u_t^\top R u_t - \gamma^2 w_t^\top w_t), \tag{5}$$

which can be solved via generalized Riccati equations [56]. 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 standard $\mathcal{H}_\infty$ control by allowing users to specify norm-bounded uncertainties on system dynamics, as seen in [57, 58].

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 defined as:

$$\arg\min_K \mathbb{E}_{\{\delta_i\},\{\gamma_i\}}[J(K,A,B)] \quad \text{s.t. } \dot{x} := \left(A + \sum_{i=1}^{p} \delta_i A_i\right) x + \left(B + \sum_{i=1}^{q} \gamma_i B_i\right) u + w,$$

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 I.

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 [59]. In this work, two approaches were proposed to learn the margin parameters, which we refer to as the "Shared Lyapunov" and "Auxiliary Stabilizer" approaches, explicitly described in their paper. While such approaches improve upon the conservatism of classical data-free margin methods by adjusting the margins to appropriately reflect the predictor error, they still require the hand specification of the perturbation matrices $\{A_i\}$ and $\{B_i\}$.

Table 1: Each of the results below are the p-values of paired t-tests conducted pairwise between methods testing $H_1 : \mathcal{R}_{\%}^{(\mathrm{CPC})} < \mathcal{R}_{\%}^{(\mathrm{alt})}$ 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**

5 Experiments

We now study five setups of interest in the infinite horizon, discrete-time, deterministic setting, namely LQR control of an airfoil [60], a load positioning system [25, 27], a Furuta pendulum [61], a DC microgrid [62], and a nuclear plant [63]. 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 nuclear plant. The full dynamics and sampling distributions of parameters are provided in Appendix J.

We compare against $\mathcal{H}_\infty$ control with γ bisection and the data-free margin methods and data-based methods for the LQR with multiplicative noise setup as discussed in Section 4. The data-free margin methods are as implemented by [54] and are fully described in Appendix I, 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 [59]. 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 [54, 55].

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. Code will be made available upon acceptance.

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 [64]. 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. $J(K^*_{\mathrm{rob}}, C) = \infty$. We, thus, only compute $\mathcal{R}_{\%}$ over the stabilizing controllers and separately report the proportion of cases for which an unstable robust controller resulted. Lower values are desirable for both.

For each comparison method, we report the result of a one-side paired t-test of $H_1 : \mathcal{R}_{\%}^{(\mathrm{CPC})} < \mathcal{R}_{\%}^{(\mathrm{alt})}$ in Table 1. We defer the presentation of the raw regret values and the percent of cases with stabilized dynamics to Appendix K due to space constraints. Notably, the alternative approaches generally incur greater regret than CPC. In the case of $\mathcal{H}_\infty$, this is expected as misspecification here is in the dynamics matrices, differing from the adversarial exogenous noise misspecification 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 3.6 and demonstrate the empirical validity of the associated assumption, we computed the empirical coverages across various levels of desired coverage $\alpha \in (0, 1)$ for the experimental setups. As previously

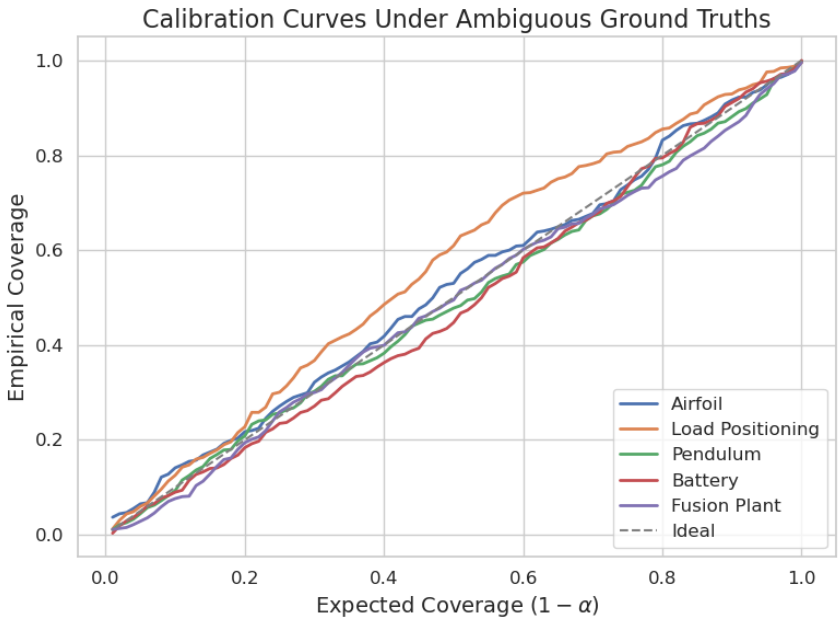

Figure 1: Calibration plots for the tasks, each assessed on 1,000 i.i.d. test samples of C with calibration performed using the estimated dynamics $\widetilde{C}$, affirming the result of Theorem 3.6.

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 robust control subproblem in a UCCD LQR 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 insights from Koopman operator theory or with nonparametric dynamics modeling using neural operators would be of interest [65, 66, 67]. Extending this to the MDP setting would also be of interest as an extension to [68].

References

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

[2] Yi Wu, Yuxin Wu, Aviv Tamar, Stuart Russell, Georgia Gkioxari, and Yuandong Tian. Learning and planning with a semantic model. *arXiv preprint arXiv:1809.10842*, 2018.

[3] Christopher T Aksland, Daniel L Clark, Christopher A Lupp, and Andrew G Alleyne. An approach to robust co-design of plant and closed-loop controller. In *2023 IEEE Conference on Control Technology and Applications (CCTA)*, pages 918–925. IEEE, 2023.

[4] Mario Garcia-Sanz. Control co-design: an engineering game changer. *Advanced Control for Applications: Engineering and Industrial Systems*, 1(1):e18, 2019.

[5] Julie A Reyer, Hosam K Fathy, Panos Y Papalambros, and A Galip Ulsoy. 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, volume 80234, pages 1023–1032. American Society of Mechanical Engineers, 2001.

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

[7] Hosam K Fathy, Julie A Reyer, Panos Y Papalambros, and AG Ulsov. On the coupling between the plant and controller optimization problems. In *Proceedings of the 2001 American Control Conference.(Cat. No. 01CH37148)*, volume 3, pages 1864–1869. IEEE, 2001.

[8] Robert Falck, Justin S Gray, Kaushik Ponnapalli, and Ted Wright. dymos: A python package for optimal control of multidisciplinary systems. *Journal of Open Source Software*, 6(59):2809, 2021.

[9] James T Allison, Tinghao Guo, and Zhi Han. Co-design of an active suspension using simultaneous dynamic optimization. *Journal of Mechanical Design*, 136(8):081003, 2014.

[10] Saeed Azad, Mohammad Behtash, Arian Houshmand, and Michael Alexander-Ramos. 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*, pages 83–97. Springer, 2018.

[11] Saeed Azad, Mohammad Behtash, Arian Houshmand, and Michael J Alexander-Ramos. Phev powertrain co-design with vehicle performance considerations using mdsdo. *Structural and Multidisciplinary Optimization*, 60:1155–1169, 2019.

[12] 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*, volume 51753, page V02AT03A003. American Society of Mechanical Engineers, 2018.

[13] 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*, volume 86229, page V03AT03A008. American Society of Mechanical Engineers, 2022.

[14] Saeed Azad and Daniel R Herber. An overview of uncertain control co-design formulations. *Journal of Mechanical Design*, 145(9):091709, 2023.

[15] Trevor J Bird, Jacob A Siefert, Herschel C Pangborn, and Neera Jain. A set-based approach for robust control co-design. *arXiv preprint arXiv:2310.11658*, 2023.

[16] Saeed Azad and Michael J Alexander-Ramos. A single-loop reliability-based mdsdo formulation for combined design and control optimization of stochastic dynamic systems. *Journal of Mechanical Design*, 143(2):021703, 2020.

[17] Tonghui Cui, Zhuoyuan Zheng, and Pingfeng Wang. Control co-design of lithium-ion batteries for enhanced fast-charging and cycle life performances. *Journal of Electrochemical Energy Conversion and Storage*, 19(3):031001, 2022.

[18] 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. *Journal of Mechanical Design*, pages 1–17, 2024.

[19] Tonghui Cui, James T Allison, and Pingfeng Wang. A comparative study of formulations and algorithms for reliability-based co-design problems. *Journal of Mechanical Design*, 142(3):031104, 2020.

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

[21] 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. *Journal of Mechanisms and Robotics*, 14(3):031016, 2022.

[22] Saeed Azad and Michael J Alexander-Ramos. Robust combined design and control optimization of hybrid-electric vehicles using mdsdo. *IEEE Transactions on Vehicular Technology*, 70(5):4139–4152, 2021.

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

[24] Saeed Azad and Michael J Alexander-Ramos. Robust mdsdo for co-design of stochastic dynamic systems. *Journal of Mechanical design*, 142(1):011403, 2020.

[25] Peyman Ahmadi, Mehdi Rahmani, and Aref Shahmansoorian. Lqr based optimal co-design for linear control systems with input and state constraints. *International Journal of Systems Science*, 54(5):1136–1149, 2023.

[26] Hosam K Fathy, Panos Y Papalambros, A Galip Ulsoy, and Davor Hrovat. Nested plant/controller optimization with application to combined passive/active automotive suspensions. In *Proceedings of the 2003 American Control Conference, 2003.*, volume 4, pages 3375–3380. IEEE, 2003.

[27] Yu Jiang, Yebin Wang, Scott A Bortoff, and Zhong-Ping Jiang. An iterative approach to the optimal co-design of linear control systems. *International Journal of Control*, 89(4):680–690, 2016.

[28] Dongdong Zhao, Xiaodi Yang, Yichang Li, Li Xu, Jinhua She, and Shi Yan. A kalman–koopman lqr control approach to robotic systems. *IEEE Transactions on Industrial Electronics*, 2024.

[29] Giorgos Mamakoukas, Maria Castano, Xiaobo Tan, and Todd Murphey. Local koopman operators for data-driven control of robotic systems. In *Robotics: science and systems*, 2019.

[30] Petar Bevanda, Max Beier, Shahab Heshmati-Alamdari, Stefan Sosnowski, and Sandra Hirche. Towards data-driven lqr with koopmanizing flows. *IFAC-PapersOnLine*, 55(15):13–18, 2022.

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

[32] Anastasios N Angelopoulos and Stephen Bates. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv preprint arXiv:2107.07511*, 2021.

[33] Glenn Shafer and Vladimir Vovk. A tutorial on conformal prediction. *Journal of Machine Learning Research*, 9(3), 2008.

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

[35] Utsav Sadana, Abhilash Chenreddy, Erick Delage, Alexandre Forel, Emma Frejinger, and Thibaut Vidal. A survey of contextual optimization methods for decision-making under uncertainty. *European Journal of Operational Research*, 2024.

[36] Abhilash Chenreddy and Erick Delage. End-to-end conditional robust optimization. *arXiv preprint arXiv:2403.04670*, 2024.

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

[38] Jan Willems. Least squares stationary optimal control and the algebraic riccati equation. *IEEE Transactions on automatic control*, 16(6):621–634, 1971.

[39] 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)*, pages 4576–4581. IEEE, 2021.

[40] Prasad Vilas Chanekar, Nikhil Chopra, and Shapour Azarm. Co-design of linear systems using generalized benders decomposition. *Automatica*, 89:180–193, 2018.

[41] Daniel R Herber and James T Allison. Nested and simultaneous solution strategies for general combined plant and control design problems. *Journal of Mechanical Design*, 141(1):011402, 2019.

[42] Abdullah Kamadan, Gullu Kiziltas, and Volkan Patoglu. Co-design strategies for optimal variable stiffness actuation. *IEEE/ASME Transactions on Mechatronics*, 22(6):2768–2779, 2017.

[43] Enrico Sisti et al. Control co-design of a co2-based chiller system. 2024.

[44] Zheng Liu, Jiaxin Wu, Wuchen Fu, Pouya Kabirzadeh, In-Bum Chung, Mohammed Jubair Dipto, Nenad Miljkovic, Pingfeng Wang, and Yumeng Li. Control co-design of battery packs with immersion cooling. In *ASME International Mechanical Engineering Congress and Exposition*, volume 87592, page V002T02A016. American Society of Mechanical Engineers, 2023.

[45] Zheng Liu, Yanwen Xu, Hao Wu, Pingfeng Wang, and Yumeng Li. Data-driven control co-design for indirect liquid cooling plate with microchannels for battery thermal management. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 87301, page V03AT03A048. American Society of Mechanical Engineers, 2023.

[46] Lennart Ljung et al. Theory for the user. *System identification*, 1987.

[47] Shai Feldman, Bat-Sheva Einbinder, Stephen Bates, Anastasios N Angelopoulos, Asaf Gendler, and Yaniv Romano. Conformal prediction is robust to dispersive label noise. In *Conformal and Probabilistic Prediction with Applications*, pages 624–626. PMLR, 2023.

[48] 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*, volume 87318, page V03BT03A035. American Society of Mechanical Engineers, 2023.

[49] Saeed Azad, Daniel R Herber, Suraj Khanal, and Gaofeng Jia. Site-dependent solutions of wave energy converter farms with surrogate models, control co-design, and layout optimization. *arXiv preprint arXiv:2405.06794*, 2024.

[50] Benjamin Gravell, Peyman Mohajerin Esfahani, and Tyler Summers. Learning optimal controllers for linear systems with multiplicative noise via policy gradient. *IEEE Transactions on Automatic Control*, 66(11):5283–5298, 2020.

[51] Maryam Fazel, Rong Ge, Sham Kakade, and Mehran Mesbahi. Global convergence of policy gradient methods for the linear quadratic regulator. In *International conference on machine learning*, pages 1467–1476. PMLR, 2018.

[52] Peter Seiler, Andrew Packard, and Pascal Gahinet. An introduction to disk margins [lecture notes]. *IEEE Control Systems Magazine*, 40(5):78–95, 2020.

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

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

[55] Benjamin Gravell, Iman Shames, and Tyler Summers. Robust data-driven output feedback control via bootstrapped multiplicative noise. In *Learning for Dynamics and Control Conference*, pages 650–662. PMLR, 2022.

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

[57] Hassan Bevrani, Mohammad Ramin Feizi, and Sirwan Ataee. Robust frequency control in an islanded microgrid: h_∞ and μ-synthesis approaches. *IEEE transactions on smart grid*, 7(2):706–717, 2015.

[58] 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. *IEEE/ASME Transactions on Mechatronics*, 20(3):1482–1490, 2014.

[59] Benjamin J Gravell, Peyman Mohajerin Esfahani, and Tyler H Summers. Robust control design for linear systems via multiplicative noise. *IFAC-PapersOnLine*, 53(2):7392–7399, 2020.

[60] Labane Chrif and Zemalache Meguenni Kadda. Aircraft control system using lqg and lqr controller with optimal estimation-kalman filter design. *Procedia Engineering*, 80:245–257, 2014.

[61] N Arulmozhi and T Victorie. Kalman filter and h∞ filter based linear quadratic regulator for furuta pendulum. *Computer Systems Science & Engineering*, 43(2), 2022.

[62] Yulin Liu, Tianhao Qie, Xinan Zhang, Hao Wang, Zhongbao Wei, Herbert HC Iu, and Tyrone Fernando. A novel online learning-based linear quadratic regulator for vanadium redox flow battery in dc microgrids. *Journal of Power Sources*, 587:233672, 2023.

[63] Hamza Boubacar Kirgni and Junling Wang. Lqr-based adaptive tsmc for nuclear reactor in load following operation. *Progress in Nuclear Energy*, 156:104560, 2023.

[64] Chunlin Sun, Linyu Liu, and Xiaocheng Li. Predict-then-calibrate: A new perspective of robust contextual lp. *arXiv preprint arXiv:2305.15686*, 2023.

[65] Steven L Brunton, Marko Budišić, Eurika Kaiser, and J Nathan Kutz. Modern koopman theory for dynamical systems. *arXiv preprint arXiv:2102.12086*, 2021.

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

[67] Elizabeth Qian, Boris Kramer, Benjamin Peherstorfer, and Karen Willcox. Lift & learn: Physics-informed machine learning for large-scale nonlinear dynamical systems. *Physica D: Nonlinear Phenomena*, 406:132401, 2020.

[68] Kai Wang, Sanket Shah, Haipeng Chen, Andrew Perrault, Finale Doshi-Velez, and Milind Tambe. Learning mdps from features: Predict-then-optimize for sequential decision making by reinforcement learning. *Advances in Neural Information Processing Systems*, 34:8795–8806, 2021.

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

[70] Jingjing Bu, Afshin Mesbahi, Maryam Fazel, and Mehran Mesbahi. Lqr through the lens of first order methods: Discrete-time case. *arXiv preprint arXiv:1907.08921*, 2019.

[71] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32, 2019.

[72] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

A Prediction Region Validity Lemma

Lemma A.1. *Let $f(K, C)$ be a function such that, for any fixed θ, it is non-negative and L-Lipschitz in $C \in \mathcal{U}(\theta)$ under the operator norm metric for any $K \in \mathcal{K}(\mathcal{U}(\theta))$, 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, for any C, that $\max_{K \in \mathcal{K}(C)} f(K, C) < \infty$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. Then, there is a function $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$ such that $\forall C \in \mathcal{C}$, $g(\{C\}, C) = 0$, $g(\Omega, C) < \infty \; \forall \Omega \in \Omega(\mathcal{C})$ such that $\operatorname{diam}(\Omega) < \infty$, and $\Omega_1 \subset \Omega_2 \implies g(\Omega_1, C) \leq g(\Omega_2, C)$ that satisfies:*

$$\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1 - \alpha. \tag{6}$$

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

$$\begin{aligned}
&\left| f(K^*(\mathcal{U}(\theta)), C) - \min_{K \in \mathcal{K}(C)} f(K, C) \right| \\
&= \left| f(K^*(\mathcal{U}(\theta)), C) - \min_{K \in \mathcal{K}(\mathcal{U}(\theta))} f(K, C) + \min_{K \in \mathcal{K}(\mathcal{U}(\theta))} f(K, C) - \min_{K \in \mathcal{K}(C)} f(K, C) \right| \\
&\leq \left| f(K^*(\mathcal{U}(\theta)), C) - \min_{K \in \mathcal{K}(\mathcal{U}(\theta))} f(K, C) \right| + \left| \min_{K \in \mathcal{K}(\mathcal{U}(\theta))} f(K, C) - \min_{K \in \mathcal{K}(C)} f(K, C) \right|
\end{aligned}$$

We now bound each term separately, starting with the first term. We first note:

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

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

$$\begin{aligned}
\left| f(K^*(\mathcal{U}(\theta)), C) - \min_{K \in \mathcal{K}(\mathcal{U}(\theta))} f(K, C) \right| &\leq \left| \min_{K \in \mathcal{K}(\mathcal{U}(\theta))} \max_{\widehat{C} \in \mathcal{U}(\theta)} f(K^*(\widehat{C}), \widehat{C}) - \min_{K \in \mathcal{K}(\mathcal{U}(\theta))} f(K, C) \right| \\
&\leq \max_{K \in \mathcal{K}(\mathcal{U}(\theta))} \left| \max_{\widehat{C} \in \mathcal{U}(\theta)} f(K, \widehat{C}) - f(K, C) \right| \\
&\leq L \max_{\widehat{C} \in \mathcal{U}(\theta)} ||\widehat{C} - C||_2 \leq L \operatorname{diam}(\mathcal{U}(\theta)).
\end{aligned}$$

The second term can be bounded simply by considering the set difference of $\mathcal{K}(C)$ and $\mathcal{K}(\mathcal{U}(\theta))$ as follows:

$$\left| \min_{K \in \mathcal{K}(\mathcal{U}(\theta))} f(K, C) - \min_{K \in \mathcal{K}(C)} f(K, C) \right| \leq \max_{K \in \mathcal{K}(C) - \mathcal{K}(\mathcal{U}(\theta))} f(K, C)$$

We then define $g(\Omega, C) := L \operatorname{diam}(\Omega) + \max_{K \in \mathcal{K}(C) - \mathcal{K}(\Omega)} f(K, C)$. To see $g(\Omega, C) < \infty \; \forall \Omega \in \Omega(\mathcal{C})$ such that $\operatorname{diam}(\Omega) < \infty$, we consider each term separately. The first follows as $\operatorname{diam}(\Omega) < \infty$. For the second, we have that:

$$\max_{K \in \mathcal{K}(C) - \mathcal{K}(\Omega)} f(K, C) \leq \max_{K \in \mathcal{K}(C)} f(K, C) < \infty,$$

where the final bound follows by assumption. To observe the singleton property, simply note that $\operatorname{diam}(\{C\}) = 0$ and $\mathcal{K}(C) - \mathcal{K}(\{C\}) = \emptyset$, meaning $\max_{K \in \mathcal{K}(C) - \mathcal{K}(\Omega)} f(K, C) = 0$, giving the desired result. To observe that g has the desired monotonocity property, we first note that, if $\Omega_1 \subset \Omega_2$, then $\operatorname{diam}(\Omega_1) \leq \operatorname{diam}(\Omega_2)$. We similarly have by assumption that, $\mathcal{K}(\Omega_2) \subset \mathcal{K}(\Omega_1)$, from which it follows that $\mathcal{K}(C) - \mathcal{K}(\Omega_1) \subset \mathcal{K}(C) - \mathcal{K}(\Omega_2)$ and immediately that the maximization problem naturally increases, namely:

$$\max_{K \in \mathcal{K}(\Omega_1) - \mathcal{K}(C)} f(K, C) \leq \max_{K \in \mathcal{K}(\Omega_2) - \mathcal{K}(C)} f(K, C).$$

g, therefore, satisfies the desired monotonicity property. Since we have the assumption that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$, the result immediately follows. □

B Deterministic Discrete-Time Regret Analysis

Theorem B.1. *Let $J(K, C)$ be the infinite horizon, discrete-time, deterministic analog of that defined in Section 2.3, i.e. $J(K, C) := \sum_{t=0}^{\infty}(x_t^\top (Q + K^\top RK)x_t)$ with $w = 0$. Assume that for any fixed θ, $K \in \mathcal{K}(\mathcal{U}(\theta))$, and $C \in \mathcal{U}(\theta)$, $D(K) := \sqrt{n}||Q + K^\top RK||_\infty ||x_0||_\infty^2 ||W||_2 < \infty$ and $||CW|| < 1$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. Then, there is a function $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$ such that $\forall C \in \mathcal{C}$, $g(\{C\}, C) = 0$, $g(\Omega, C) < \infty \; \forall \Omega \in \Omega(\mathcal{C})$ such that $\mathrm{diam}(\Omega) < \infty$, and $\Omega_1 \subset \Omega_2 \implies g(\Omega_1, C) \leq g(\Omega_2, C)$ that satisfies:*

$$\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1 - \alpha. \tag{7}$$

Proof. We consider any fixed θ and demonstrate that $J(K, C)$ is non-negative and L-Lipschitz in $C \in \mathcal{U}(\theta)$ under the operator norm metric for any $K \in \mathcal{K}(\mathcal{U}(\theta))$, from which Lemma A.1 can be invoked to arrive at the desired conclusion. Given the assumed determinism of the dynamics, we have that $x_t = (CW)^t x_0$, meaning the above objective setup can equivalently be expressed as:

$$\sum_{t=0}^{\infty} x_0^\top ((CW)^{t\top}(Q + K^\top RK)(CW)^t)x_0, \tag{8}$$

J is clearly non-negative by construction. To further see that, for any C, $\max_{K \in \mathcal{K}(C)} J(K, C) < \infty$, we notice that:

$$\begin{aligned}\left|\left|\sum_{t=0}^{\infty} x_0^\top ((CW)^{t\top}(Q + K^\top RK)(CW)^t)x_0\right|\right|_2 &\leq \sum_{t=0}^{\infty} \left|\left|x_0^\top ((CW)^{t\top}(Q + K^\top RK)(CW)^t)x_0\right|\right|_2 \\ &\leq \sum_{t=0}^{\infty} ||x_0||_2^2 ||Q + K^\top RK||_2 ||CW||_2^{2t} = \frac{||x_0||_2^2 ||Q + K^\top RK||_2}{1 - ||CW||^2} < \infty,\end{aligned}$$

where the finiteness of the final quantity is guaranteed by the assumptions on $D(K)$ and $||CW|| < 1$.

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 C, which we do as follows

$$\begin{aligned}&\nabla_C \left(\sum_{t=0}^{\infty} x_0^\top ((CW)^{t\top}(Q + K^\top RK)(CW)^t)x_0\right) \\ &= \sum_{t=0}^{\infty} t\mathrm{diag}((Q + K^\top RK)(CW)^t x_0)(CW)^{(t-1)}\mathrm{diag}(x_0)W^\top \\ &+ \sum_{t=0}^{\infty} t\mathrm{diag}((Q^\top + (RK)^\top K)(CW)^t x_0)(CW)^{(t-1)}\mathrm{diag}(x_0)W^\top\end{aligned}$$

We now bound the magnitude of this quantity as follows:

$$\begin{aligned}L &\leq \max_C ||\sum_{t=0}^{\infty} t\mathrm{diag}((Q + K^\top RK)(CW)^t x_0)(CW)^{(t-1)}\mathrm{diag}(x_0)W^\top \\ &+ \sum_{t=0}^{\infty} t\mathrm{diag}((Q^\top + (RK)^\top K)(CW)^t x_0)(CW)^{(t-1)}\mathrm{diag}(x_0)W^\top||_2 \\ &\leq \max_C \sum_{t=0}^{\infty} t\left|\left|\mathrm{diag}((Q + K^\top RK)(CW)^t x_0)(CW)^{(t-1)}\mathrm{diag}(x_0)W^\top\right|\right|_2 \\ &+ \sum_{t=0}^{\infty} t\left|\left|\mathrm{diag}((Q^\top + (RK)^\top K)(CW)^t x_0)(CW)^{(t-1)}\mathrm{diag}(x_0)W^\top\right|\right|_2,\end{aligned}$$

where we have used $\mathrm{diag}(x_0)$ 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 $||\mathrm{diag}(x_0)||_2 = ||x_0||_\infty$ and $||A||_\infty \leq \sqrt{n}||A||_2$ for $A \in \mathbb{R}^{n \times n}$ as follows:

$$\max_C \sum_{t=0}^{\infty} t||\text{diag}((Q + K^\top RK)(CW)^t x_0)(CW)^{(t-1)}\text{diag}(x_0)W^\top||_2$$

$$\leq \max_C \sum_{t=0}^{\infty} t||\text{diag}((Q + K^\top RK)(CW)^t x_0)||_2||(CW)^{(t-1)}||_2||\text{diag}(x_0)||_2||W^\top||_2$$

$$= \max_C \sum_{t=0}^{\infty} t||(Q + K^\top RK)(CW)^t x_0||_\infty||x_0||_\infty||(CW)^{(t-1)}||_2||W||_2$$

$$\leq \max_C \sum_{t=0}^{\infty} t||Q + K^\top RK||_\infty||(CW)^t||_\infty||x_0||_\infty||x_0||_\infty||(CW)^{(t-1)}||_2||W||_2$$

$$\leq \max_C \sum_{t=0}^{\infty} t(\sqrt{n}||Q + K^\top RK||_\infty||x_0||_\infty^2||W||_2)||(CW)^t||_2||(CW)^{(t-1)}||_2.$$

We now collect all terms independent of t into a constant $D(K) = \sqrt{n}||Q + K^\top RK||_\infty||x_0||_\infty^2||W||_2$, leaving us with:

$$\leq \max_C D(K) \sum_{t=0}^{\infty} t\,||CW||_2^{2t-1} = \max_C D(K) \frac{||CW||_2}{\left(||CW||_2^2 - 1\right)^2}$$

which converges if $||CW||_2 < 1$, as assumed. □

The finite LQR case follows immediately as a corollary of the above, stated below for completeness. Notably, the qualitative nature of the two differs in that no assumptions on the underlying dynamics are necessary in the finite time horizon case, i.e. $||CW||_2 < 1$ is no longer a necessary assumption, as the bound of the suboptimality will necessarily be finite simply by virtue of considering a finite time horizon.

Corollary B.2. *Let $J(K, C)$ be the finite horizon, discrete-time, deterministic analog of that defined in Section 2.3, i.e. $J(K, C) := \sum_{t=0}^{T}(x_t^\top(Q + K^\top RK)x_t)$ with $w = 0$. Assume that for any fixed θ, $K \in \mathcal{K}(\mathcal{U}(\theta))$, and $C \in \mathcal{U}(\theta)$, $D(K) := \sqrt{n}||Q + K^\top RK||_\infty||x_0||_\infty^2||W||_2 < \infty$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. Then, there is a function $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$ such that $\forall C \in \mathcal{C}$, $g(\{C\}, C) = 0$, $g(\Omega, C) < \infty\ \forall \Omega \in \Omega(\mathcal{C})$ such that $\text{diam}(\Omega) < \infty$, and $\Omega_1 \subset \Omega_2 \implies g(\Omega_1, C) \leq g(\Omega_2, C)$ that satisfies:*

$$\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1 - \alpha. \tag{9}$$

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. *Let $J(K, C)$ be the infinite horizon, continuous-time, deterministic analog of that defined in Section 2.3, i.e. $J(K, C) := \int_0^\infty (x(t)^\top(Q + K^\top RK)x(t))dt$ for $w = 0$. Assume that for any fixed θ, $K \in \mathcal{K}(\mathcal{U}(\theta))$, and $C \in \mathcal{U}(\theta)$, $D(K) := \sqrt{n}||Q + K^\top RK||_\infty||x_0||_\infty^2||W||_2 < \infty$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. Then, there is a function $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$ such that $\forall C \in \mathcal{C}$, $g(\{C\}, C) = 0$, $g(\Omega, C) < \infty\ \forall \Omega \in \Omega(\mathcal{C})$ such that $\text{diam}(\Omega) < \infty$, and $\Omega_1 \subset \Omega_2 \implies g(\Omega_1, C) \leq g(\Omega_2, C)$ that satisfies:*

$$\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1 - \alpha. \tag{10}$$

Proof. We consider any fixed θ and demonstrate the desired properties that $J(K, C)$ is non-negative and L-Lipschitz in $C \in \mathcal{U}(\theta)$ under the operator norm metric for any $K \in \mathcal{K}(\mathcal{U}(\theta))$, from which Lemma A.1 can be invoked to arrive at the desired conclusion. Given the assumed determinism of the dynamics, we further have that $x(t) = e^{CWt}x(0)$, meaning the above objective setup can equivalently be expressed with the uncertainty sets related to the objective function, namely as:

$$\int_0^\infty x(0)^\top(e^{CWt})^\top(Q + K^\top RK)(e^{CWt})x(0)dt. \tag{11}$$

J is clearly non-negative by construction. To further see that, for any C, $\max_{K\in\mathcal{K}(C)} J(K,C) < \infty$, we notice that:

$$\begin{aligned}
&\left|\left|\int_0^\infty (x(0)^\top (e^{CWt})^\top (Q+K^\top RK)(e^{CWt})x(0))dt\right|\right|_2 \\
&\leq \int_0^\infty \left|\left|x(0)^\top (e^{CWt})^\top (Q+K^\top RK)(e^{CWt})x(0)\right|\right|_2 dt \\
\leq \int_0^\infty ||x(0)||_2^2 ||e^{CWt}||_2^2 ||Q+K^\top RK||_2^2 dt &= ||x(0)||_2^2 ||Q+K^\top RK||_2^2 \int_0^\infty ||e^{CWt}||_2^2 dt \\
\leq ||x(0)||_2^2 ||Q+K^\top RK||_2^2 \int_0^\infty \beta(C)^2 e^{-2\alpha(C)t} dt &= \frac{\beta(C)^2 ||x(0)||_2^2 ||Q+K^\top RK||_2^2}{2\alpha(C)} < \infty
\end{aligned}$$

where we used the fact for any $K \in \mathcal{K}(C)$, we have that CW is Hurwitz by definition, which exhibits the standard matrix exponential bound $||e^{CWt}|| \leq \beta(C)e^{-\alpha(C)t}$ for some $\alpha(C), \beta(C) > 0$.

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 as follows:

$$\begin{aligned}
&\nabla_C \left(\int_0^\infty x(0)^\top (e^{CWt})^\top (Q+K^\top RK)(e^{CWt})x(0)dt\right) \\
&= \int_0^\infty t\text{diag}((Q+K^\top RK)e^{CWt}x(0))e^{CWt}\text{diag}(x(0))W^\top dt \\
&+ \int_0^\infty t\text{diag}((Q^\top + (RK)^\top K)e^{CWt}x(0))e^{CWt}\text{diag}(x(0))W^\top dt
\end{aligned}$$

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

$$\begin{aligned}
&\max_C \left|\left|\int_0^\infty t\text{diag}((Q+K^\top RK)e^{CWt}x(0))e^{CWt}\text{diag}(x(0))W^\top dt\right|\right|_2 \\
&\leq \max_C \int_0^\infty t \left|\left|\text{diag}((Q+K^\top RK)e^{CWt}x(0))\right|\right|_2 \left|\left|e^{CWt}\right|\right|_2 ||\text{diag}(x(0))||_2 \left|\left|W^\top\right|\right|_2 dt \\
&= \max_C \int_0^\infty t \left|\left|(Q+K^\top RK)e^{CWt}x(0)\right|\right|_\infty \left|\left|e^{CWt}\right|\right|_2 ||x(0)||_\infty ||W||_2 dt \\
&\leq \max_C \int_0^\infty t(\sqrt{n}\left|\left|Q+K^\top RK\right|\right|_\infty ||W||_2 ||x(0)||_\infty^2) \left|\left|e^{CWt}\right|\right|_2^2 dt.
\end{aligned}$$

Collecting all terms independent of t into a constant $D(K) = \sqrt{n}||Q+K^\top RK||_\infty ||W||_2 ||x(0)||_\infty^2$ and using the bound $||e^{CWt}|| \leq \beta(C)e^{-\alpha(C)t}$, we reach the conclusion as:

$$= \max_C D(K) \int_0^\infty t \left|\left|e^{CWt}\right|\right|_2^2 dt \leq D(K)\beta(C)^2 \int_0^\infty te^{-2\alpha(C)t} dt = \max_C \frac{D(K)\beta(C)^2}{4\alpha(C)^2},$$

as desired. □

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

Corollary C.2. *Let $J(K,C)$ be the finite horizon, continuous-time, deterministic analog of that defined in Section 2.3, i.e. $J(K,C) := \int_0^T (x(t)^\top (Q+K^\top RK)x(t))dt$ for $w = 0$. Assume that for any fixed θ, $K \in \mathcal{K}(\mathcal{U}(\theta))$, and $C \in \mathcal{U}(\theta)$, $D(K) := \sqrt{n}||Q+K^\top RK||_\infty ||x_0||_\infty^2 ||W||_2 < \infty$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1-\alpha$. Then, there is a function $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$ such that $\forall C \in \mathcal{C}$, $g(\{C\}, C) = 0$, $g(\Omega, C) < \infty\ \forall \Omega \in \Omega(\mathcal{C})$ such that $\text{diam}(\Omega) < \infty$, and $\Omega_1 \subset \Omega_2 \implies g(\Omega_1, C) \leq g(\Omega_2, C)$ that satisfies:*

$$\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1-\alpha. \tag{12}$$

D Stochastic Discrete-Time Regret Analysis

Theorem D.1. *Let $J(K,C)$ be the infinite horizon, discrete-time, stochastic analog of that defined in Section 2.3, i.e. $J(K,C) := \sum_{t=0}^\infty (x_t^\top (Q+K^\top RK)x_t)$ with $w_t \sim \mathcal{N}(0,\Sigma)$ i.i.d. across t. Assume that for any fixed θ, $K \in \mathcal{K}(\mathcal{U}(\theta))$,*

and $C \in \mathcal{U}(\theta)$: (1) $||Q_t + K^\top R_t K|| \leq \beta_1 e^{-\alpha_1 t}$ for some constants $\alpha_1, \beta_1 > 0$ such that $\min_C(2\alpha_2(C) + \alpha_1) > 0$ for $\alpha_2(C) := \min_{K \in \mathcal{K}(C)}(-\log(||CW||))$ and (2) $D_1(K) := \max_t \sqrt{n}||Q(t) + K^\top R(t)K||_\infty ||x_0||_\infty^2 ||W||_2 < \infty$, $D_2(K) := ||\Sigma||_2 ||W||_2 < \infty$, and $||CW|| < 1$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. Then, there is a function $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$ such that $\forall C \in \mathcal{C}$, $g(\{C\}, C) = 0$, $g(\Omega, C) < \infty \; \forall \Omega \in \Omega(\mathcal{C})$ such that $\operatorname{diam}(\Omega) < \infty$, and $\Omega_1 \subset \Omega_2 \implies g(\Omega_1, C) \leq g(\Omega_2, C)$ that satisfies:

$$\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1 - \alpha. \tag{13}$$

Proof. We again consider any fixed θ and demonstrate the desired properties that $J(K, C)$ is non-negative and L-Lipschitz in $C \in \mathcal{U}(\theta)$ under the operator norm metric for any $K \in \mathcal{K}(\mathcal{U}(\theta))$, from which Lemma A.1 can be invoked to arrive at the desired conclusion. Notice the objective can be reformulated in the standard manner as follows:

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

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

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

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

With these simplifications, we are left with:

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

J is clearly non-negative by construction. To further see that, for any C, $\max_{K \in \mathcal{K}(C)} J(K, C) < \infty$, we first note that the first term was bounded in the proof of Theorem B.1, where the assumptions on $D_1(K)$ and $||CW||$ permit its invocation. For the second term, we proceed similarly, as follows:

$$\begin{aligned}
&\leq \sum_{t=0}^{\infty}\sum_{k=0}^{t-1}\left|\left|\operatorname{Tr}((Q_t + K^\top R_t K)(CW)^k \Sigma (CW)^{k\top}))\right|\right|_2 \\
&\leq ||\Sigma||_2 \sum_{t=0}^{\infty}||Q_t + K^\top R_t K||_2 \sum_{k=0}^{t-1}||CW||_2^{2k} \leq ||\Sigma||_2 \int_0^{\infty}||Q_t + K^\top R_t K||_2 \int_0^t ||CW||_2^{2k} dk dt \\
&\leq \frac{||\Sigma||_2}{2\log(||CW||_2)}\int_0^{\infty}||Q_t + K^\top R_t K||_2 (||CW||_2^{2t} - 1)dt \\
&\leq \beta_1 \frac{||\Sigma||_2}{2\log(||CW||_2)}\int_0^{\infty} e^{-\alpha_1 t}(||CW||_2^{2t} - 1)dt < \infty
\end{aligned}$$

where we used the assumption that $||CW||_2 < 1$ to ensure convergence of the integral.

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:

$$\sum_{t=0}^{\infty}\sum_{k=0}^{t-1}\nabla_C \mathrm{Tr}((Q_t + K^\top R_t K)(CW)^k \Sigma (CW)^{k\top}))$$
$$= \sum_{t=0}^{\infty}\sum_{k=0}^{t-1} k((((Q_t + K^\top R_t K)^\top (CW)^k \Sigma^\top) \odot (CW)^{(k-1)})W^\top +$$
$$\sum_{t=0}^{\infty}\sum_{k=0}^{t-1} k(((Q_t + K^\top R_t K)(CW)^k \Sigma) \odot (CW)^{(k-1)})W^\top$$

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.

$$\max_C || \sum_{t=0}^{\infty}\sum_{k=0}^{t-1} k((((Q_t + K^\top R_t K)^\top (CW)^k \Sigma^\top) \odot (CW)^{(k-1)})W^\top ||_2$$
$$\leq \max_C \sum_{t=0}^{\infty}\sum_{k=0}^{t-1} k||((Q_t + K^\top R_t K)^\top (CW)^k \Sigma^\top) \odot (CW)^{(k-1)}||_2 ||W||_2$$
$$\leq \max_C \sum_{t=0}^{\infty}\sum_{k=0}^{t-1} k||Q_t + K^\top R_t K||_2 ||(CW)^k||_2 ||\Sigma||_2 ||(CW)^{(k-1)}||_2 ||W||_2$$
$$\leq \max_C \sum_{t=0}^{\infty}\sum_{k=0}^{t-1} k||Q_t + K^\top R_t K||_2 ||\Sigma||_2 ||W||_2 ||CW||_2^{(2k-1)}$$

We again now collect all terms independent of t into a constant $D_2(K) = ||\Sigma||_2||W||_2$. We prove the convergence by first expressing $\alpha_2(C) \leq -\log(||CW||) \implies ||CW|| \leq e^{-\alpha_2(C)}$, giving us

$$\leq \max_C D_2(K)\beta_1 \int_0^\infty e^{-\alpha_1 t} \int_0^t k e^{-2\alpha_2(C)k} dk dt$$
$$\leq \max_C \frac{D_2(K)\beta_1}{4\alpha_2(C)^2} \int_0^\infty e^{-\alpha_1 t} \left(e^{-2\alpha_2(C)t}(-2\alpha_2(C)t + e^{2\alpha_2(C)t} - 1) \right) dt$$
$$= \max_C \frac{D_2(K)\beta_1}{\alpha_1(2\alpha_2(C) + \alpha_1)^2},$$

where the final convergence is guaranteed if $\min_C(2\alpha_2(C) + \alpha_1) > 0$ as desired. □

E Stochastic Continuous-Time Regret Analysis

Theorem E.1. *Let $J(K, C)$ be the infinite horizon, continuous-time, stochastic analog of that defined in Section 2.3, i.e. $J(K, C) := \mathbb{E}\left[\int_0^\infty (x(t)^\top (Q(t) + K^\top R(t) K) x(t)) dt\right]$ with $w(t)$ a white noise process with spectral density Σ. Assume that for any fixed θ, $K \in \mathcal{K}(\mathcal{U}(\theta))$, and $C \in \mathcal{U}(\theta)$: (1) there exist constants $\alpha_1, \beta_1 > 0$ such that $||Q(t) + K^\top R(t) K|| \leq \beta_1 e^{-\alpha_1 t}$ and $\min_C(2\alpha_2(C) + \alpha_1) > 0$ where $\alpha_2(C) := \max_{K \in \mathcal{K}(C)}(-\max_i \mathrm{Re}(\lambda_i(CW)))$ and (2) $D_1(K) := \max_t \sqrt{n}||Q(t) + K^\top R(t) K||_\infty ||x_0||_\infty^2 ||W||_2 < \infty$ and $D_2(K) := ||\Sigma||_2 ||W||_2 < \infty$. Assume further that $\mathcal{P}_{\Theta,C}(C \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. Then, there is a function $g : \Omega(\mathcal{C}) \times \mathcal{C} \to \mathbb{R}$ such that $\forall C \in \mathcal{C}$, $g(\{C\}, C) = 0$, $g(\Omega, C) < \infty \; \forall \Omega \in \Omega(\mathcal{C})$ such that $\mathrm{diam}(\Omega) < \infty$, and $\Omega_1 \subset \Omega_2 \implies g(\Omega_1, C) \leq g(\Omega_2, C)$ that satisfies:*

$$\mathcal{P}_{\Theta,C}\left(0 \leq \mathcal{R}(\Theta, C) \leq g(\mathcal{U}(\Theta), C)\right) \geq 1 - \alpha. \tag{14}$$

Proof. We again consider any fixed θ and demonstrate the desired properties that $J(K, C)$ is non-negative and L-Lipschitz in $C \in \mathcal{U}(\theta)$ under the operator norm metric for any $K \in \mathcal{K}(\mathcal{U}(\theta))$, from which Lemma A.1 can be invoked to arrive at the desired conclusion. Notice the objective can be reformulated in the standard manner as follows:

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

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 [69]:

$$J(K,C) = \int_0^\infty x(0)^\top (e^{CWt})^\top (Q(t) + K^\top R(t)K)(e^{CWt})x(0)dt$$
$$+ \int_0^\infty \int_0^t \mathrm{Tr}((Q(t) + K^\top R(t)K)e^{CWk}\Sigma e^{CWk^\top})dkdt$$

J is clearly non-negative by construction. To further see that, for any C, $\max_{K\in\mathcal{K}(C)} J(K,C) < \infty$, we first note that the first term was bounded in the proof of Theorem C.1. For the second term, we proceed similarly, as follows:

$$\left|\left|\int_0^\infty \int_0^t \mathrm{Tr}((Q(t) + K^\top R(t)K)e^{CWk}\Sigma e^{CWk^\top})dkdt\right|\right|_2$$
$$\leq \int_0^\infty \int_0^t \left|\left|\mathrm{Tr}((Q(t) + K^\top R(t)K)e^{CWk}\Sigma e^{CWk^\top})\right|\right|_2 dkdt$$
$$\leq \int_0^\infty \int_0^t ||\Sigma||_2||Q(t) + K^\top R(t)K||_2||e^{CWk}||_2^2 dkdt$$
$$\leq ||\Sigma||_2 \int_0^\infty \beta_1 e^{-\alpha_1 t}\int_0^t \beta_2(C)^2 e^{-2\alpha_2(C)k}dkdt$$
$$= ||\Sigma||_2 \frac{\beta_1\beta_2(C)^2}{2\alpha_2(C)}\int_0^\infty e^{-\alpha_1 t}\left(1 - e^{-2\alpha_2(C)t}\right)dt = ||\Sigma||_2 \frac{\beta_1\beta_2(C)^2}{2\alpha_2(C)}\left(\frac{1}{\alpha_1} - \frac{1}{\alpha_1 + 2\alpha_2(C)}\right)$$
$$= ||\Sigma||_2 \frac{\beta_1\beta_2(C)^2}{\alpha_1(\alpha_1 + 2\alpha_2(C))} < \infty$$

where we used the assumption that $\min_C(\alpha_1 + 2\alpha_2(C)) > 0$ to ensure convergence and fact that for any $K \in \mathcal{K}(C)$, CW is Hurwitz by definition, which exhibits the well-known matrix exponential bound $||e^{CWt}|| \leq \beta_2(C)e^{-\alpha_2(C)t}$.

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 $D_1(K)$. We, thus, solely present that of the second term as follows:

$$\int_0^\infty \int_0^t \nabla_C \mathrm{Tr}((Q(t) + K^\top R(t)K)e^{(CW)k}\Sigma e^{(CW)k^\top})dkdt$$
$$= \int_0^\infty \int_0^t k(((Q(t)^\top + (R(t)K)^\top K)e^{kCW}\Sigma^\top) \odot e^{kCW})W^\top dkdt$$
$$+ \int_0^\infty \int_0^t k(((Q(t) + K^\top R(t)K)e^{kCW}\Sigma) \odot e^{kCW})W^\top$$

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.

$$L \leq \max_C ||\int_0^\infty \int_0^t k(((Q(t)^\top + (R(t)K)^\top K)e^{kCW}\Sigma^\top) \odot e^{kCW})W^\top dkdt||_2$$
$$\leq \max_C \int_0^\infty \int_0^t k||Q(t) + K^\top R(t)K||_2||\Sigma||_2||W||_2||e^{kCW}||_2^2 dkdt$$

We again now collect all terms independent of t into a constant $D_2(K) = ||\Sigma||_2||W||_2$, leaving

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

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

F Coverage Guarantees Under Noisy Observations

Theorem F.1. *Let $\widetilde{C} = C + \epsilon$ where vec$(\epsilon) \sim \mathcal{N}(0, \Sigma)$. Assume $\mathcal{U}(\theta) = \{C' \mid ||f(\theta) - C'||_2 \leq \widehat{q}\}$ satisfies $\mathcal{P}_{\Theta,\widetilde{C}}(\widetilde{C} \in \mathcal{U}(\Theta)) \geq 1 - \alpha$. If for any $\theta \in \Theta$ and $\delta > 0$, $\mathcal{P}_{C|\Theta=\theta}(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2) > \mathcal{P}_{C|\Theta=\theta}(\widehat{q}^2 \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2 + \delta)$, then*

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

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

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

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

$$\begin{aligned}
&:= \mathcal{P}_{C,\delta|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 + \delta \leq \widehat{q}^2\right) \\
&= \mathcal{P}_{C|\Theta=\theta,\delta>0}\left(\|C - f(\theta)\|_2^2 + \delta \leq \widehat{q}^2\right)\mathcal{P}_{C|\Theta=\theta}(\delta > 0) \\
&\quad + \mathcal{P}_{C|\Theta=\theta,\delta\leq 0}\left(\|C - f(\theta)\|_2^2 + \delta \leq \widehat{q}^2\right)\mathcal{P}_{C|\Theta=\theta}(\delta \leq 0) \\
&= \mathcal{P}_{C|\Theta=\theta,\delta>0}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2 - \delta\right)\mathcal{P}_{C|\Theta=\theta}(\delta > 0) \\
&\quad + \mathcal{P}_{C|\Theta=\theta,\delta>0}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2 + \delta\right)\mathcal{P}_{C|\Theta=\theta}(\delta \leq 0).
\end{aligned}$$

$$
\begin{aligned}
&= \mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right) \\
&\quad - \mathcal{P}_{C|\Theta=\theta}\left(\delta > 0\right)\mathcal{P}_{C|\Theta=\theta,\delta>0}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right) \\
&\quad - \mathcal{P}_{C|\Theta=\theta}\left(\delta \leq 0\right)\mathcal{P}_{C|\Theta=\theta,\delta\leq 0}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right) \\
&\quad + \mathcal{P}_{C|\Theta=\theta}\left(\delta > 0\right)\mathbb{E}_{\delta>0}\left[\mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2 - \delta\right)\right] \\
&\quad + \mathcal{P}_{C|\Theta=\theta}\left(\delta \leq 0\right)\mathbb{E}_{\delta>0}\left[\mathcal{P}_{C,\delta|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2 + \delta\right)\right] \\
&= \mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right) \\
&\quad - \mathcal{P}_{C|\Theta=\theta}\left(\delta > 0\right)\mathbb{E}_{\delta>0}\left[\mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right) - \mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2 - \delta\right)\right] \\
&\quad + \mathcal{P}_{C|\Theta=\theta}\left(\delta \leq 0\right)\mathbb{E}_{\delta>0}\left[\mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2 + \delta\right) - \mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right)\right] \\
&= \mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right) \\
&\quad - \mathcal{P}_{C|\Theta=\theta}\left(\delta > 0\right)\mathbb{E}_{\delta>0}\left[\mathcal{P}_{C|\Theta=\theta}\left(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right)\right] \\
&\quad + \mathcal{P}_{C|\Theta=\theta}\left(\delta \leq 0\right)\mathbb{E}_{\delta>0}\left[\mathcal{P}_{C|\Theta=\theta}\left(\widehat{q}^2 \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2 + \delta\right)\right].
\end{aligned}
$$

We, therefore, have that $\mathcal{P}_{\widetilde{C}|\Theta=\theta}(\widetilde{C} \in \mathcal{U}(\theta)) \leq \mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right) + \Delta$, where

$$
\begin{aligned}
\Delta := &\mathcal{P}_{C|\Theta=\theta}\left(\delta \leq 0\right)\mathbb{E}_{\delta>0}\left[\mathcal{P}_{C|\Theta=\theta}\left(\widehat{q}^2 \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2 + \delta\right)\right] \\
&- \mathcal{P}_{C|\Theta=\theta}\left(\delta > 0\right)\mathbb{E}_{\delta>0}\left[\mathcal{P}_{C|\Theta=\theta}\left(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right)\right]
\end{aligned}
$$

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

$$
\mathcal{P}_{C|\Theta=\theta}\left(\widehat{q}^2 - \delta \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2\right) > \mathcal{P}_{C|\Theta=\theta}\left(\widehat{q}^2 \leq \|C - f(\theta)\|_2^2 \leq \widehat{q}^2 + \delta\right).
$$

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

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

Therefore, $\mathcal{P}_{\widetilde{C}|\Theta=\theta}\left(\left\|\widetilde{C} - f(\theta)\right\|_2 \leq \widehat{q}\right) - \mathcal{P}_{C|\Theta=\theta}\left(\|C - f(\theta)\|_2 \leq \widehat{q}\right) \leq \Delta \leq 0$ □

G LQR C Gradient

We follow the presentation of [51] 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 G.1. *Let $J(K, C)$ be the infinite horizon, discrete-time, deterministic analog of that defined in Section 2.3, i.e. $J(K, C) := \sum_{t=0}^{\infty}(x_t^\top(Q + K^\top RK)x_t)$ for $w = 0$. Then,*

$$
\nabla_C J(K, C) = 2P_K CWX_K W^T, \tag{15}
$$

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

Proof. By the standard reformulation of $J(K, C)$ as described in [51], 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

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

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

H Policy Gradient Convergence Guarantee

Lemma H.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:*

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

Then $\phi(x) := \max_{y \in \mathcal{Y}} f(x,y)$ is c^*-gradient dominated, where $c^* := \max_{y \in \mathcal{Y}} c(y)$.

Proof. The proof for this follows immediately from Danskin's Theorem, which states that $\nabla \phi(x) = \nabla_x f(x, y^*(x))$ for $y^*(x) := \max_{y \in \mathcal{Y}} f(x,y)$, meaning:

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

□

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 H.1, from which we reach the desired conclusion. The former fact was demonstrated in [70], which we present below for sake of convenience with modification of notational conventions to match that used herein.

Lemma H.2. *(Lemma 3 of [51]) Let $J(K,C)$ be the infinite horizon, discrete-time, deterministic analog of that defined in Section 2.3, i.e. $J(K,C) := \sum_{t=0}^{\infty} (x_t^\top (Q + K^\top R K) x_t)$ for $w = 0$. Then, if $X_K \succcurlyeq 0$ and $K \in \mathcal{K}(C)$,*

$$J(K,C) - J(K^*(C),C) \leq \frac{||X_{K^*(C)}||}{\sigma_{\min}(X_K)^2 \sigma_{\min}(R)} ||\nabla_K J(K,C)||_F^2, \tag{16}$$

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

Lemma H.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 Section 2.3, i.e. $J(K,C) := \sum_{t=0}^{\infty} (x_t^\top (Q + K^\top R K) x_t)$ for $w = 0$. Then, for $K \in \mathcal{K}(\mathcal{C})$ where $X_K \succcurlyeq 0$ for all $\widehat{C} \in \mathcal{C}$, $\phi(K)$ is $\mu(K)$-gradient dominated for $\mu(K) := \max_{\widehat{C} \in \mathcal{C}} \frac{||X_{K^*(\widehat{C})}||}{\sigma_{\min}(X_K)^2 \sigma_{\min}(R)}$, where X_K and P_K respectively solve the following two Lyapunov equations: $\Delta_K X_K \Delta_K^\top - X_K = 0$ and $P_K = \Delta_K^\top P_K \Delta_K + Q + K^\top R K$, where $\Delta_K := A - BK$. That is, $\phi(K) - \phi(K_{\text{rob}}^*(\mathcal{C})) \leq \mu(K) ||\nabla_K \phi(K)||_F^2$.*

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

Theorem H.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 Section 2.3, i.e. $J(K,C) := \sum_{t=0}^{\infty}(x_t^\top(Q+K^\top RK)x_t)$ 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*

$$\eta_K \leq \min_{[\widehat{A},\widehat{B}]:=\widehat{C}\in\mathcal{C}} \frac{1}{16}\min\Bigg\{\left(\frac{\sigma_{\min}(Q)\nu}{J(K,\widehat{C})}\right)^2 \frac{1}{||\widehat{B}||||\nabla_K J(K,\widehat{C})||(1+||\widehat{A}-\widehat{B}K||)}, \frac{\sigma_{\min}(Q)}{2J(K,\widehat{C})||R+\widehat{B}^\top P_K\widehat{B}||}\Bigg\}. \tag{17}$$

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 [51], specifically in their presentation of Lemma 24, in which we leverage the above developed gradient dominance result, namely that in Lemma H.3. We first note that Algorithm 1 is equivalent to performing gradient descent over $\phi(K)$ upon assuming convergence of the inner maximization over C, that is if $C^{(T_C)} = C^*(K^{(t)})$. It, therefore, suffices to characterize gradient descent, i.e. $K^{(t+1)} := K^{(t)} - \eta_K\nabla_K\phi(K^{(t)})$.

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

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

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

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

where we take $\gamma := \min_t \mu(K^{(t)})\gamma(K^{(t)})$. We now turn to proving $\phi(K^{(t)})-\phi(K^{(t+1)}) \geq \gamma(K^{(t)})||\nabla_K\phi(K^{(t)})||$. To do so, we use an intermediate result in the proof of Lemma 24 in [51], in which it was demonstrated that for any fixed dynamics C, there is a $\beta(C)>0$ such that $J(K,C)-J(K',C)\geq\beta(C)||\nabla_K J(K,C)||$ 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 (17). To leverage this result, we must, therefore, re-express the quantity of interest into an expression with fixed dynamics:

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

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

I 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 Section 4, 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

$$J(A, B, \{A_i\}, \{B_i\}, \{\delta_i\}, \{\gamma_i\}, K) := \int_0^\infty (x^\top Q x + (Kx)^\top R(Kx))dt \tag{18}$$
$$\text{s.t.} \quad \dot{x} = \left((A + \sum_{i=1}^p \delta_i A_i) - (B + \sum_{i=1}^q \gamma_i B_i)K\right) x.$$

The problem specifications, therefore, vary in the K used as the stopping criterion of Equation (18) 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 [71] with optimization done using Adam [72] 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.

J 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.

J.1 Aircraft Control

We consider the experimental setup studied in [60], 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	$\gamma_\beta, \gamma_p, \gamma_r, \gamma_{\delta_r}, \gamma_{\delta_a}$	$\mathcal{N}(\mu_\gamma, \Sigma_\gamma)$	$\mu_\gamma \sim \mathcal{U}([0,1]^5), \Sigma_\gamma = AA^\top,\ A \sim \mathcal{U}([0,1]^{5\times 5})$
L coefficients	$L_\beta, L_p, L_r, L_{\delta_r}, L_{\delta_a}$	$\mathcal{N}(\mu_L, \Sigma_L)$	$\mu_L \sim \mathcal{U}([0,1]^5), \Sigma_L = AA^\top,\ A \sim \mathcal{U}([0,1]^{5\times 5})$
N coefficients	$N_\beta, N_p, N_r, N_{\delta_r}, N_{\delta_a}$	$\mathcal{N}(\mu_N, \Sigma_N)$	$\mu_N \sim \mathcal{U}([0,1]^5), \Sigma_N = AA^\top,\ A \sim \mathcal{U}([0,1]^{5\times 5})$

J.2 Load Positioning Control

We consider the load-positioning system of [25, 27]. 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	$d_B = u \cdot m_B$	$u \sim \mathcal{U}(0.004,\ 0.0667)$

J.3 Furuta Pendulum

We also consider the Furuta pendulum dynamical system given in [61], in which the system dynamics were specified by

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

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

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	$\lvert\mathcal{N}(\mu_{D_r}, \sigma^2_{D_r})\rvert$	$\mu_{D_r} = 0.0015,\ \sigma_{D_r} \sim \mathcal{U}(0,1)$

J.4 DC Microgrids

We additionally consider the LQR model of DC microgrids given in [62], in which the system dynamics were specified by

$$A = \begin{bmatrix} \frac{2(-u_0 d - NK_2S)}{V_s d} & 0 & \frac{-2NK_4S}{V_s d} & \frac{-4NK_5S}{V_s d} & \frac{2u_0}{V_s} & 0 & 0 & 0 & 0 \\ 0 & \frac{2(-u_0 d - NK_3S)}{V_s d} & \frac{4NK_4S}{V_s d} & \frac{6NK_5S}{V_s d} & 0 & \frac{2u_0}{V_s} & 0 & 0 & 0 \\ \frac{6NK_2S}{V_s d} & \frac{4NK_3S}{V_s d} & \frac{2(-u_0 d - NK_4S)}{V_s d} & 0 & 0 & 0 & \frac{2u_0}{V_s} & 0 & 0 \\ \frac{-4NK_2S}{V_s d} & \frac{-2NK_3S}{V_s d} & 0 & \frac{2(-u_0 d - NK_5S)}{V_s d} & 0 & 0 & 0 & \frac{2u_0}{V_s} & 0 \\ \frac{u_0}{V_t} & 0 & 0 & 0 & \frac{-u_0}{V_t} & 0 & 0 & 0 & 0 \\ 0 & \frac{u_0}{V_t} & 0 & 0 & 0 & \frac{-u_0}{V_t} & 0 & 0 & 0 \\ 0 & 0 & \frac{u_0}{V_t} & 0 & & 0 & \frac{-u_0}{V_t} & 0 & 0 \\ 0 & 0 & 0 & \frac{u_0}{V_t} & 0 & 0 & 0 & \frac{-u_0}{V_t} & 0 \\ \frac{NRT}{FC_2^c} & \frac{-NRT}{FC_3^c} & \frac{NRT}{FC_4^c} & \frac{NRT}{FC_5^c} & 0 & 0 & 0 & 0 & 0 \end{bmatrix} B = \begin{bmatrix} \frac{C_2^t - C_2^c}{V_s/2} \\ \frac{C_3^t - C_3^c}{V_s/2} \\ \frac{C_4^t - C_4^c}{V_s/2} \\ \frac{C_5^t - C_5^c}{V_s/2} \\ \frac{C_2^c - C_2^t}{V_t} \\ \frac{C_3^c - C_3^t}{V_t} \\ \frac{C_4^c - C_4^t}{V_t} \\ \frac{C_5^c - C_5^t}{V_t} \\ 0 \end{bmatrix}$$

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

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	$\mathcal{N}(\mu_{C_5^t}, \sigma^2_{C_5^t})$	$\mu_{C_5^t} = 1.0,\ \sigma_{C_5^t} = 0.667$

J.5 Nuclear Plant

We finally consider the terminal sliding-mode control of a nuclear plant from [63], 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 nuclear 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	$\mathcal{N}(\mu_{X_0}, \sigma^2_{X_0})$	$\mu = 1.0,\ \sigma = 1.0$

K Additional Experimental Results

We here provide the raw regret results corresponding to the experiment shown in Section 5.1 in Table 7 and the proportion of stabilized 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: Each of the results below are 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