# Robust Linear Quadratic Regulator: Exact Tractable Reformulation

Wouter Jongeneel, Tyler Summers and Peyman Mohajerin Esfahani

***Abstract*— We consider the problem of controlling an unknown stochastic linear dynamical system subject to an infinite-horizon discounted quadratic cost. Existing approaches for handling the corresponding robust optimal control problem resort to either conservative uncertainty sets or various approximations schemes, and to our best knowledge, the current literature lacks an exact, yet tractable, solution. We propose a class of novel uncertainty sets for the system matrices of the linear system. We show that the resulting robust linear quadratic regulator problem enjoys a closed-form solution described through a generalized algebraic Riccati equation arising from dynamic game theory.**

## I. INTRODUCTION

A broad variety of problems from engineering, machine learning, and operations research involve optimizing the behaviour of a dynamical system in the face of inherent uncertainties in the system model used for design and decision-making. A vast literature going back several decades has studied various aspects of this robust control problem, including substantial work on system identification; adaptive, robust, and optimal control, *e.g.,* see [1]–[3].

In this work we consider the discrete-time Linear Quadratic Regulator (LQR) problem under parametric uncertainties. Ever since the LQR problem originated, robustness was questioned. It is known that the discrete-time LQR can suffer from the lack of a stability margin [4], or if any, it is typically a noticeably worse margin in comparison with the continuous-time counterpart [5]. Moreover, our understanding of the corresponding perturbation theory is limited [6], [7]. The inherent presence of uncertainties in practice indeed reinforces the need to address these issues. A classical $\mu$-synthesis approach is generally intractable [8], [9] while a tractable LMI approach like proposed in [10] may be conservative. This work investigates to what extend dynamic game theory can be a middle-ground.

### A. Related Work

This paper is centered around quantifying the robustness resulting from a dynamic game with quadratic cost and linear dynamics. Early accounts of this viewpoint can be found on for example page 90 of the monograph by Whittle [11]. There, the remark is made that extremizing a *risk-sensitive* multi-stage optimal control cost function can be interpreted as another, yet now *constrained*, optimal control problem.

The authors are with the Delft Center for Systems and Control, TU Delft, The Netherlands (P.MohajerinEsfahani@tudelft.nl, W.Jongeneel@student.tudelft.nl) and the Control, Optimization, and Networks lab, UT Dallas, The United States (tyler.summers@utdallas.edu). The work of T. Summers was sponsored by the Army Research Office and was accomplished under Grant Number: W911NF-17-1-0058.

There is a large body of work in this direction. The celebrated paper [12] provides necessary and sufficient conditions for the continuous-time system $\dot{x}(t) = (A + \Delta_A(t))x(t) + (B + \Delta_B(t))u(t)$, $(\Delta_A \quad \Delta_B) = DF(t)(E_1 \quad E_2)$, $\|F(t)\| \leq 1$ to be stabilizable. This result was later generalized to the discrete-time case in [13]. Although these results are more than 20 years old, describing parametric uncertainties in the pair $(A, B)$ via some matrix-norm-balls is still the prevalent method, however currently driven by measure concentration results, *e.g.,* see [14], [15]. In the stochastic case, distributional uncertainties in the form of *relative entropy* constraints are considered [16], [17].

Although these problems are well understood, the catch within this game theoretic framework is that, the uncertainty set typically depends on the extremizing parameters. Therefore, it is not clear, *a priori*, over which set of models the robust control problem is solved, this is effectively only known *a posteriori*. Moreover, most results do not consider the full uncertainty set their optimization problem can handle, but rather focus on some "*inscribed ball*", see [17, ch 10] on how to fit an ellipsoid to data. Motivated by renewed interest in tractable reformulations of (Robust) LQR problems (*cf.* [18]–[23]), we investigate which lessons can be drawn from the readily available dynamic game theory.

### B. Contribution and Outline

This work focuses on a novel formulation and solution of a robust LQR problem. Our contributions are as follows:

(i) We propose a novel family of uncertainty sets for the system matrices, and show that the worst-case cost over these sets can be solved efficiently (Proposition III.6).
(ii) Given the proposed uncertainty sets, we develop an exact, up to an algebraic Riccati equation, solution to the corresponding Robust LQR problem (Theorem III.7).

The article is structured as follows. In Section II, we formally introduce several key definitions along with the robust optimal control problem that will be addressed. The new uncertainty set and the corresponding main results are presented in Section III, which are interpreted from a game theoretic point of view in Section IV. In Section V, we illustrate the presented results through a numerical example.

Due to the lack of space, we refer the interested readers to an extended version of this work [24]. This version contains the technical proofs along with new results and discussions, including an extension to uncertain input matrices and further structural properties of our uncertainty set and worst-case models.

***Notation:*** We use standard notation, but to be clear. Let $\mathbb{R}_{\geq 0}$ denote the set of non-negative real numbers, whereas $I_n$ is the identity element of $\mathbb{R}^{n\times n}$. Let $\mathcal{S}_+^n$ be the cone of symmetric positive semi-definite matrices on which the ordering is denoted by $A \succeq B$. The largest singular-value of a matrix $A$ equals $\|A\|_2$. Let $\mathrm{Tr}(\cdot)$ be the trace operator, then the inner-product between $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times n}$ is defined as $\langle A, B\rangle = \mathrm{Tr}(A^\top B)$ such that $\langle A, A\rangle = \|A\|_F^2$ for $\|\cdot\|_F$ the Frobenius-norm. Similarly, $\|X\|_{F,Q}^2$ is used to denote $\mathrm{Tr}(X^\top Q X)$ for $Q \succ 0$. Furthermore, when $A$ is said to be *exponentially stable* its spectrum is fully contained in the open unit disk. The expectation operator is given by $\mathbb{E}[\cdot]$ and $X \sim \mathcal{P}(\mu, \Sigma)$ is a random variable with mean $\mu$ and covariance $\Sigma$ for a distribution $\mathcal{P}$. Optimality is indicated with a $\star$, so $x^\star$ is for example the minimizer of a function $f(x)$ with $f^\star = f(x^\star)$. Also, in the context of an optimization program, *s.t.* stands for *subject to*.

## II. PRELIMINARIES

In this section the problem at hand is introduced.

### A. Robust LQR problem

Given the matrices $Q \in \mathcal{S}_+^n, R \in \mathcal{S}_{++}^m$, discount factor $\alpha \in (0,1)$ and $\widehat{A} \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$, $\Sigma_0, \Sigma_v \in \mathcal{S}_{++}^n$, and $\{v_k\}_{k\in\mathbb{N}}$ being a white noise sequence of independent random variables with zero mean and a time-invariant covariance matrix $\Sigma_v$, *i.e.*, $\mathbb{E}[v_i] = 0$ and $\mathbb{E}[v_i v_j^\top] = \delta_{ij}\Sigma_v$ for all $i, j \in \mathbb{N}$. Then we seek an optimal policy $\pi^\star = \{\mu_0^\star, \mu_1^\star, \dots\}$ that solves the discounted *Robust Linear-Quadratic Regulator* (RLQR) problem over the uncertainty set $\mathbb{A}$:

$$\begin{aligned} \inf_{\{\mu_k\}_{k=0}^\infty} \sup_{\Delta_A} \quad & \mathop{\mathbb{E}}_{x_0, v}\left[\sum_{k=0}^\infty \alpha^k \left(x_k^\top Q x_k + u_k^\top R u_k\right)\right], \\ \text{s.t.} \quad & x_{k+1} = (\widehat{A} + \Delta_A)x_k + Bu_k + v_k, \\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0, \Sigma_v), \quad x_0 \sim \mathcal{P}(0, \Sigma_0), \\ & u_k = \mu_k(x_k), \quad \Delta_A \in \mathbb{A}. \end{aligned} \tag{1}$$

In other words, we consider the LQR problem where the system matrix $A$ is not precisely known, but known to be described by $A = \widehat{A} + \Delta_A$. Here our prior estimate of $A$ is denoted by $\widehat{A}$, whereas $\Delta_A \in \mathbb{A}$ is the uncertainty. A particular example of such a setting naturally emerges in statistics or identification problems where $\widehat{A}$ is the current estimate of $A$ and $\mathbb{A}$ contains $\Delta_A$ with high probability.

See [24] for an extension to the case where also $B$ is partially unknown.

**Assumption II.1** (Linear time-invariant policy)**:** In problem (1), we restrict the class of control policies $\mu_k$ to linear time-invariant (LTI) controllers $\mu_k(x) = Kx$ where $K \in \mathbb{R}^{m\times n}$.

Instead of writing the full program (1) over again, introduce a compact notation:

**Definition II.2** (Discounted LQ cost)**:** Consider the dynamical system $x_{k+1} = Ax_k + v_k$ where the noise process and the intial condition follow $v_k \overset{i.i.d.}{\sim} \mathcal{P}(0, \Sigma_v)$ and $x_0 \sim \mathcal{P}(0, \Sigma_0)$. Then we define the linear quadratic (LQ) cost function $\mathcal{J} : \mathbb{R}^{n\times n} \times \mathcal{S}_+^n \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ by

$$\mathcal{J}(A, Q) := \mathop{\mathbb{E}}_{x_0, v}\left[\sum_{k=0}^\infty \alpha^k x_k^\top Q x_k\right].$$

Since we consider a *discounted* LQ cost, it is helpful to also introduce a respective notion of stability.

**Definition II.3** ($\sqrt{\alpha}$-stability)**:** Let $\alpha \in (0,1]$, then the matrix $A$ is $\sqrt{\alpha}$-stable when its spectrum is fully contained in the open disk with radius $\alpha^{-1/2}$, *i.e.,* $\sqrt{\alpha}A$ is exponentially stable.

One can observe that the classical exponential stability notion in system theory is a sufficient condition, and *not* necessary, for the $\sqrt{\alpha}$-stability of Definition II.3.

The main objective of this study is to introduce an uncertainty set $\mathbb{A}$ that facilitates an exact and tractable robust LQR formulation which is meaningful to study. To this end, we first proceed with a brief discussion regarding a desirable property of such an uncertainty set $\mathbb{A}$.

### B. Convexity in Robust Linear Control

As the next example shows Assumption II.1 restricts possible $\mathbb{A}$. There is no time-invariant $K$ which can stabilize all stabilizable pairs $(A, B)$:

**Example II.4** (Lack of universal stabilizing feedback law)**:** Consider for some finite scalar $c$ and $d \in (-1, 1)$ the matrices

$$A_1 = \begin{pmatrix} 1 & c \\ 0 & d \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & c \\ 0 & d \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The pairs $(A_1, B)$ and $(A_2, B)$ are stabilizable. However if we let the controller be of the form $K = \begin{pmatrix} K_1 & K_2 \end{pmatrix}$ then $(A_1, B)$ needs $K_1 \in (-2, 0)$ while $(A_2, B)$ needs $K_1 \in (0, 2)$ to make the closed-loop matrix exponentially stable. Since $(-2, 0) \cap (0, 2) = \{\emptyset\}$ there is no $K$ which can exponentially stabilize both systems.

Example II.4 can be interpreted in the spirit of *switching control*, *i.e.,* once $(A_1, B)$ switches to system $(A_2, B)$ your linear control law should switch as well. As indicated by the discrete-time version of Lemma 3.1 from [25], given a compact subset $\mathcal{K} \subset \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$ of the set of stabilizable pairs $(A, B)$ one can indeed introduce a finite covering where all the elements of each segment can be stabilized via a common feedback gain, *e.g.,* $(A_1, B)$ and $(A_2, B)$ are never members of the same segment while for example $\|A_1\|_2 = \|A_2\|_2$. This simple observation indicates that the *existence* of a stabilizing solution to (1) is not immediately obvious, even for simple norm-balls.

One may wonder how these individual segments look like, and in particular with the desire of a tractable algorithm in mind, whether a set of stabilizable pairs with a common stabilizing feedback is necessarily convex in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$. The following example provides a negative answer to this question.

**Example II.5** (Non-convex segment)**:** Consider for $a = 2$ and $d = 0.5$ the matrices

$$A_1 = \begin{pmatrix} d & 0 & a \\ 0 & d & 0 \\ 0 & 0 & d \end{pmatrix}, \quad A_2 = A_1^\top, \quad B = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Then $(A_1, B)$ and $(A_2, B)$ are both stabilizable, perhaps by $K = d^2 B^\top$, while for $A = 0.5A_1 + 0.5A_2$ the pair $(A, B)$ is not stabilizable. Moreover, since one can find a path from $(A_1, B)$ to $(A_2, B)$ which can be stabilized by $K$, there does exist some non-convex segment containing both of the pairs.

These quick examples indicate that convex uncertainty sets for $(A, B)$ in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$, which currently dominate the field, can be a restrictive point of view indeed and may be potentially conservative. Now, note that we do not claim that non-convexity is desirable, but merely observe that it should not be ruled out.

## III. MAIN RESULTS

The main objective of this section is to provide a closed-form solution to the RLQR problem (3) and study its implications.

### A. Introduction of a new uncertainty set

**Definition III.1** (Uncertainty set)**:** Given a tuple $(\widehat{A}, D, \Sigma_0, \Sigma_v, \alpha)$ and some $\gamma \in \mathbb{R}_{\geq 0}$, let $W_{0,v} := \Sigma_0 + \alpha(1-\alpha)^{-1}\Sigma_v$ and define a set of models around $\widehat{A}$ by the set:

$$\mathcal{A}_\gamma(\widehat{A}) := \left\{ A \in \mathbb{R}^{n\times n} \; : \; \begin{array}{l} A = \widehat{A} + D\Delta_A, \\ \Sigma_x = \alpha A \Sigma_x A^\top + W_{0,v}, \\ \Sigma_x \succ 0, \\ \langle \Delta_A^\top \Delta_A, \Sigma_x \rangle \leq \gamma \end{array} \right\}. \tag{2}$$

For notational convenience, we shall refer to the collection of $\Delta_A$ by $\mathbb{\Delta}_\gamma(\widehat{A})$. Using this notation, we therefore have the following simple relation between these sets: $\mathcal{A}_\gamma(\widehat{A}) = \widehat{A} + D\mathbb{\Delta}_\gamma(\widehat{A})$[1].

**Remark III.2** (Absence of translation invariance)**:** Let $B_r(x)$ be an Euclidean ball with radius $r$ and center $x$. Then one can think of $\mathcal{A}_\gamma(\widehat{A})$ as a ball with radius $\gamma$ and center $\widehat{A}$. However, in contrast to an Euclidean ball, our set is not translation invariant and depends on the center $\widehat{A}$. Moreover, since $W_{0,v} \succ 0$, for $\Delta_A$ to be in $\mathbb{\Delta}_\gamma(\widehat{A})$ is the same as being part of the set $\{\Delta_A \in \mathbb{R}^{d\times n} \; : \; \|\Delta_A^\top\|^2_{F,\Sigma_x} \leq \gamma\}$ for $\Sigma_x$ as in (2). This further explains why $\gamma$ is referred to as a "*radius*".

**Remark III.3** (Structural information)**:** The matrix $D$ in Definition III.1 may be used to incorporate a form of prior structural information into the uncertainty set. Without any prior structural information, one should choose $D = I_n$.

Before addressing (1) under (2), we provide, inspired by Lemma 2 from [22], some insights about the set $\mathcal{A}_\gamma$, which are especially interesting from an optimization point of view

[1] With slight abuse of notation, by $+$ between two sets we mean the Minkowski sum: $A + B = \{a + b \; : \; a \in A, b \in B\}$.

**Proposition III.4:** The set $\mathcal{A}_\gamma(\widehat{A})$ as defined in Definition III.1 has the following properties:

(i) For $n \geq 3$ there are sets $\mathcal{A}_\gamma(\widehat{A})$ which are non-convex.

(ii) For $\gamma > 0$, the set $\mathcal{A}_\gamma(\widehat{A})$ is semi-algebraic.

Further extending the tools from [22] to the game theoretic regime, allows for showing that the set is path-connected. The fact that our uncertainty set is semi-algebraic and does not rule out the lack of convexity is nice from a control theoretic point of view as well. See for example [26], an Euclidean ball of $(A, B) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$ intersected with the set of controllable pairs $(A, B)$ is semi-algebraic. In the following we illustrate the general non-convexity of the proposed uncertainty sets through an example.

**Example III.5** (Non-convexity of $\mathbb{\Delta}_\gamma$)**:** We consider the case where the uncertainty set $\mathcal{A}_\gamma \subset \mathbb{R}^{3\times 3}$ from Definition III.1 is constructed using the parameters $\alpha = 0.95$, $D = Q_{c\ell} = \Sigma_0 = I_3$, and $\Sigma_v = 0.01I_3$. Here we will consider several "levels sets" of $\mathbb{\Delta}_\gamma$. Since the set $\mathbb{\Delta}_\gamma$ is essentially a 9-dimensional object, for the sake of illustration we restrict our attention to a 2-dimensional subset. For this purpose, we consider the closed loop matrix $A_{c\ell}$, and especially all $\Delta_A$, to be parametrized by

$$A_{c\ell} = \underbrace{\begin{pmatrix} 0.25 & 1.25 & -0.84 \\ 0 & 0.25 & 0 \\ 0.70 & 1.25 & 0.25 \end{pmatrix}}_{\widehat{A}+BK} + \underbrace{\begin{pmatrix} 0 & 0 & \Delta_{A13} \\ 0 & 0 & 0 \\ \Delta_{A31} & 0 & 0 \end{pmatrix}}_{\Delta_A},$$

where $\Delta_{A13} = 4.98\theta_1 - 0.25\theta_2$, $\Delta_{A31} = 0.45\theta_2 - 1.08\theta_1$, and the parameters $(\theta_1, \theta_2)$ belong to the interval $[-1, 1]^2$. This choice of $(\theta_1, \theta_2)$ over $(\Delta_{A13}, \Delta_{A31})$ is purely driven by visualization purposes. Figure 1a depicts the 2-dimensional slice of $\mathbb{\Delta}_\gamma$ by means of $(\theta_1, \theta_2)$ for the levels: $\gamma \in \{2^{-4}, 2^1, 2^4, 2^{14}\}$. Interestingly enough, it is non-convex for large values of $\gamma$. Figure 1b also illustrates the LQ cost $\mathcal{J}(A_{c\ell}, Q_{c\ell})$ from Definition II.2.

At last, using the shorthand notation, the problem (1) over (2) is written as

$$\inf_{K\in\mathbb{R}^{n\times m}} \sup_{A_{c\ell}\in\mathcal{A}_\gamma(\widehat{A}+BK)} \mathcal{J}(A_{c\ell}, Q + K^\top RK). \tag{3}$$

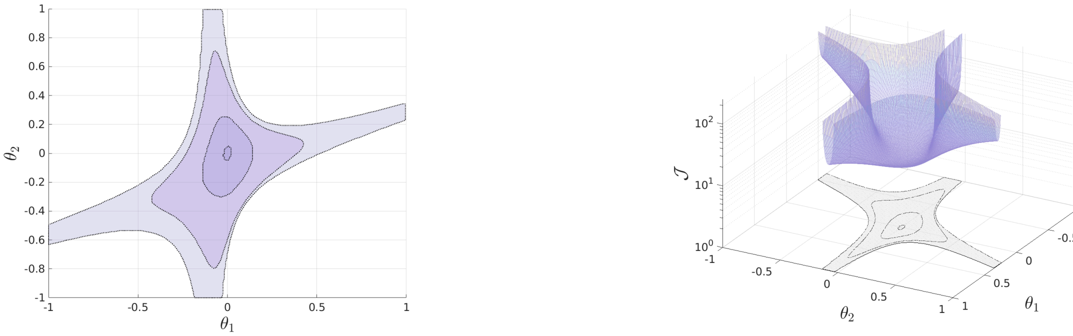
It is worth noting the dependence on $K$ in the inner maximization step. A solution to (3) is given by $\big(K^\star(\gamma), A^\star_{c\ell}(\gamma)\big)$.

### B. Solving a Robust LQR Problem

In the first step, we tackle the worst-case LQ problem over $\mathcal{A}_\gamma$, being the inner maximization of the RLQR problem (3). This problem is defined as

$$\sup_{A_{c\ell}\in\mathcal{A}_\gamma(\widehat{A}_{cl})} \mathcal{J}(A_{c\ell}, Q_{c\ell}), \tag{4}$$

for some given controller $K$ $\sqrt{\alpha}$-stabilizing $\widehat{A}_{c\ell} := \widehat{A} + BK$ and $Q_{c\ell} := Q + K^\top RK$ being the closed-loop cost matrix. Denote the solution to (4) by by $A^\star_{c\ell}(\gamma) := \widehat{A}_{c\ell} + D\Delta^\star_A(\gamma)$.




(a) Uncertainty set $\mathbb{A}_\gamma$ as in Definition III.1.

(b) LQ cost $\mathcal{J}$ as in Definition II.2.

Fig. 1: Given the parameters from Example III.5 we show the uncertainty sets and LQ function parametrized by $(\theta_1, \theta_2) \in [-1, 1]^2$ for different levels $\gamma \in \{2^{-4}, 2^1, 2^4, 2^{14}\}$ from darker to lighter gray.

**Proposition III.6** (Worst-case LQ cost)**:** Consider problem (4) with nominal closed-loop model $\widehat{A}_{c\ell}$, structural matrix $D$, some $\alpha \in (0, 1)$, initial data $\Sigma_0, \Sigma_v \in \mathcal{S}^n_{++}$, and closed-loop cost matrix $Q_{c\ell} \in \mathcal{S}^n_+$. Given some $\delta \in \mathbb{R}_{\geq 0}$, let us assume that $(\delta^{-1} I_d - \alpha D^\top S D) \succ 0$ is satisfied for the (minimal) positive semi-definite solution $S$ to the algebraic equation

$$S = Q_{c\ell} + \alpha \widehat{A}_{\mathrm{cl}}^\top \big(S + \alpha S D (\delta^{-1} I_d - \alpha D^\top S D)^{-1} D^\top S\big) \widehat{A}_{\mathrm{cl}}.$$

Then define

$$\Delta^\star_A(\delta) = (\delta^{-1} I_d - \alpha D^\top S D)^{-1} D^\top S \widehat{A}_{\mathrm{cl}}.$$

Further, define $\widetilde{\Sigma}_x$ as the positive-definite solution to the Lyapunov equation

$$\widetilde{\Sigma}_x = \alpha \big(\widehat{A}_{c\ell} + D\Delta^\star_A(\delta)\big) \widetilde{\Sigma}_x \big(\widehat{A}_{c\ell} + D\Delta^\star_A(\delta)\big)^\top + W_{0,v} \tag{5}$$

which in its turn defines the function

$$\widetilde{h}(\delta) = \left\langle \big(\Delta^\star_A(\delta)\big)^\top \Delta^\star_A(\delta), \widetilde{\Sigma}_x \right\rangle. \tag{6}$$

Then, $\Delta^\star_A(\gamma) = \Delta^\star_A(\delta)$ and $\mathcal{J}^\star = \langle \widetilde{\Sigma}_x, Q_{c\ell} \rangle$ are the optimizer (worst-case uncertainty) and the optimal value of the problem (4) with the parameter $\gamma = \widetilde{h}(\delta)$.

Now we are at the stage to address (3). This is not completely new, see for example [16], [27], where in the former[2], the pair $(\gamma, \delta)$ is interpreted via multiplier theory. We provide, in line with Definition III.1, a slightly different *system*- instead of *signal*- theoretic interpretation.

**Theorem III.7** (Optimal Robust LQ regulator)**:** Consider the RLQR problem (3) with the nominal $\sqrt{\alpha}$-stabilizable model $(\widehat{A}, B)$, the structural matrix $D$, $\alpha \in (0, 1)$, the cost matrices $Q \in \mathcal{S}^n_+, R \in \mathcal{S}^m_{++}$ and the covariance matrices $\Sigma_v, \Sigma_0 \in \mathcal{S}^n_{++}$. Given the parameter $\delta \in \mathbb{R}_{\geq 0}$, assume that the algebraic equation

$$P = Q + \alpha \widehat{A}^\top P \big(I_n + \alpha(BR^{-1}B^\top - \delta DD^\top)P\big)^{-1} \widehat{A} \tag{7}$$

in $P$ admits a minimal[3] positive semi-definite solution denoted $P(\delta)$ and define $\Lambda(\delta)$ correspondingly via $\Lambda := I_n + \alpha(BR^{-1}B^\top - \delta DD^\top P$. Furthermore, define

$$\Delta^\star_A(\delta) = \alpha\delta D^\top P(\delta) \big(\Lambda(\delta)\big)^{-1} \widehat{A} \tag{8}$$

and let $\widehat{A}^\star_{c\ell}(\gamma) := \widehat{A} + D\Delta^\star_A(\delta) + BK^\star(\gamma)$. Next, consider the expressions for $\widetilde{\Sigma}_x$ and $\widetilde{h}(\delta)$ as in (5) and (6) respectively, which are now functions of $K$ as well, to emphasize the difference, the tildes are dropped, *i.e.,* define:

$$\Sigma_x = \alpha \widehat{A}^\star_{c\ell}(\gamma) \Sigma_x \big(\widehat{A}^\star_{c\ell}(\gamma)\big)^\top + W_{0,v} \tag{9}$$

$$h(\delta) = \left\langle \big(\Delta^\star_A(\delta)\big)^\top \Delta^\star_A(\delta), \Sigma_x \right\rangle. \tag{10}$$

Then,

(i) the controller $u_k = K^\star(\gamma) x_k$ defined by

$$K^\star(\gamma) = -\alpha R^{-1} B^\top P(\delta) \big(\Lambda(\delta)\big)^{-1} \widehat{A} \tag{11}$$

is (the minimizing part of) the solution to the RLQR problem for $\gamma = h(\delta)$.

(ii) Furthermore, the maximizing solution is $\widehat{A}^\star_{c\ell}(\gamma)$, differently put, the worst-case[4] system matrix is given by $A^\star(\gamma) = \widehat{A} + D\Delta^\star_A(\delta)$.

(iii) At last, the map $h(\delta)$ is analytic and non-decreasing over some interval $[0, \overline{\delta}) \subset \mathbb{R}_{\geq 0}$ for $\overline{\delta} < \infty$.

See section IV for a game theoretic interpretation of this "*breakdown point*" $\overline{\delta}$.

It is also important to remark that although problem (1) is well-defined for all $\gamma \in \mathbb{R}_{\geq 0}$, Theorem III.7 does not simply hold for any $\gamma \in \mathbb{R}_{\geq 0}$ but rather for some range $[0, \overline{\gamma}) \subseteq \mathbb{R}_{\geq 0}$ where $h(\overline{\delta}) = \overline{\gamma}$. See Section 5.2.2 from [24] for a discussion on the properties of this map $h$, we do not necessarily have $\lim_{\delta \uparrow \overline{\delta}} h(\delta) = \infty$. This explains the implicit formulation of the Theorem.

Additionally, despite this work being about $A$, we can make a remark regarding $B$. As shown in [13], [29], when $\det(\widehat{A}) \neq 0$ then, theoretically, the extension to case with a partially known $B$ is available by simply extending the state

[2]Specifically, see sec. 2.4 and ch.7-8 for a discussion.

[3]See chapter 3 from [28] for the definition and more information.

[4]In the appendix of [24] we affirmatively answer the question if this worst-case model is actually a least-favourable model.

space and applying the aforementioned theory. See [24] for a further discussion and more ideas.

Finally, recall that Theorem III.7 presents us with an explicit expression for the worst-case model. From there we can infer further structural properties which is discussed at length in [24]. These observations are interesting since game theoretic formulations play a prominent role, either explicitly or implicitly, in many control-related fields. Also, it is expected that these results are quite general since they hinge on the symmetries in the cost. So, for the better or worse, even the most basic game theoretic robust control formulation displays a rich structure. It is the authors hope that this inspires further investigations in tractable robust control algorithms while alleviating predominant conservatism.

## IV. Game Theoretic Interpretation of Robustness

In this section we sketch the proof of the main results via a brief discussion on the connection of the original problem to dynamic game theory setting. We note that we are not the first to spot this link between game- and control theory, see for example [28] and references therein. Given the parameters $Q \in \mathcal{S}^n_+$, $R \in \mathcal{S}^m_{++}$, and $\delta \in [0, \overline{\delta}) \subseteq \mathbb{R}_{\geq 0}$, we define the function $g : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$ by

$$g(x,u,w) = \left(x^\top Q x + u^\top R u - \delta^{-1} w^\top w\right)$$

and consider for some $\alpha \in (0,1)$ the *stochastic (discounted) two-player zero-sum dynamic game* defined as:

$$\begin{aligned} \inf_{\{\mu_k\}_{k=0}^\infty} \sup_{\{\nu_k\}_{k=0}^\infty} \quad & \mathbb{E}_{x_0,v}\left[\sum_{k=0}^\infty \alpha^k g(x_k,u_k,w_k)\right], \\ \text{s.t.} \quad & x_{k+1} = Ax_k + Bu_k + Dw_k + v_k, \\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0,\Sigma_v), \quad x_0 \sim \mathcal{P}(0,\Sigma_0), \\ & u_k = \mu_k(x_k), \quad w_k = \nu_k(x_k). \end{aligned} \tag{12}$$

Here, the parameter $\delta$ penalizes the input of the $\nu$-player, whose objective it is to destabilize the system, see [28] for conditions under which (12) can be solved. Note that this game is "*diagonal*" in the sense that there are no cross-terms in the cost. This form is chosen to keep the exposition simple, but one can consider more involved adversarial terms, *e.g.,* $w_k^\top S w_k$ for some $S \succeq 0$. Nevertheless, this program heavily relies on the single parameter $\delta$. The parameter $\delta$ is constrained to live in the interval $[0,\overline{\delta})$, where $\overline{\delta}$ is referred to as the *breakdown point*, beyond this value, the $\nu$-player has to pay so little that the it can steer the cost to infinity[5].

To see a relationship between dynamic game theory and parametric uncertainty sets, suppose (12) admits a solution, then consider the following. The policy of the $\nu$-player aims at maximizing the cost. But since the $\mu$-player can handle this *worst-case* policy, it must also be able to handle policies of a *less* powerful adversary. This effectively gives rise to a whole *family* of state feedback policies the $\mu$-player can handle.

[5]See ch.8 [16] for more on the relation between this breakdown point and $\mathcal{H}_\infty$ control.

Using the Lagrangian formulation from constrained optimization, one can take the adversarial part out of the cost and put it into the constraints, *i.e.,* let $g$ be redefined as $g(x,u) = x^\top Q x + u^\top R u$ and add a constraint of the form $\mathbb{E}_{x_0,v}\left[\alpha^k w_k^\top w_k\right] \leq \gamma$, for some $\gamma \in \mathbb{R}_{\geq 0}$. Then, Theorem III.7 establishes a link between $\gamma$ and $\delta$ of the form $h(\delta) = \gamma$ such that we can relate their solutions as well.

## V. Numerical Example

The goal of this section is to compare the actions of a nominal control law to our robust framework. Consider the controllable pair $(\widehat{A}, B)$ and the structural matrix $D$ defined as

$$\widehat{A} = \begin{pmatrix} 1.2 & 0.5 \\ 0 & 1.2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Also define the covariance matrices $\Sigma_v = 0.1 I_2$, $\Sigma_0 = I_2$, the cost matrices $Q = 0.1 I_2$, $R = 10$, and the discount factor $\alpha = 0.95$.

Then, set $K$ to the nominal discounted LQ regulator[6], *i.e.,* $K = K^\star(0)$. Now, Figure 2a depicts the level sets of $\mathbb{A}_\gamma\big(\widehat{A} + BK^\star(0)\big)$ as defined by Definition III.1 for different levels[7] $\gamma \in \Gamma := \{0.005, 0.03, 0.09, 0.4, 1\}$. We further solve the worst-case model uncertainty problem (4) via Proposition III.6. Let us recall that the mapping $\widetilde{h}$ defined in (6) provides the relation $\gamma = \widetilde{h}(\delta)$ between the different values of $\gamma$. In this example, the corresponding $\delta$ are $10^{-3} \cdot \{2, 3.9, 5.5, 7.3, 7.7\}$. The locations of these worst-case models are marked by a star symbol in Figure 2a.

Looking at the cost in Figure 2a anyone could guess where these worst-case uncertainties reside. However, we have only considered this low-dimensional example for visualization purposes. Computationally speaking, nothing prohibits us from doing high dimensional examples (*e.g.,* $n = 1000$), and then Proposition III.6 might help in indicating where your system is sensitive with respect to the cost.

Finally, it is interesting to highlight what a robust controller $K^\star(\gamma)$ would do, for say, $\gamma_5 = 1$. See Figure 2b for the corresponding sets $\mathbb{A}_{\gamma=1}$ under both types of controller. When compared with Figure 2a, we see that the robust controller anticipates on where the troubles might occur.

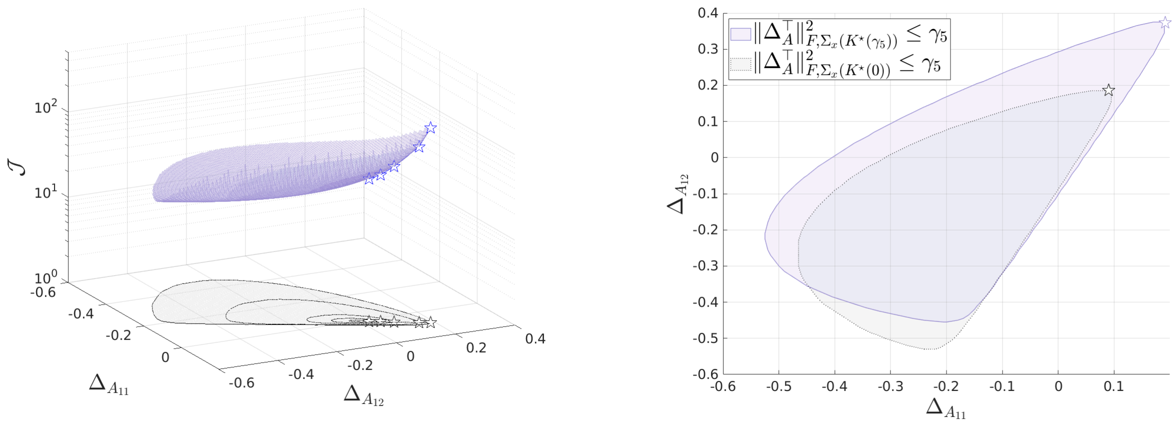
**Remark V.1** (From radius to feedback)**:** In [24] we provide tools to do the aforementioned computations efficiently, which hinge on Theorem III.7.(iii). For example, let us be given a desired "*radius*" $\gamma$ and assume it is feasible in the sense of Theorem III.7. Moreover, let the (local) Lipschitz constant of the map $h$ be some $L \in \mathbb{R}_{>0}$ on $[0,\delta)$ and select $\beta \in \mathbb{R}_{>0} : \beta \leq L^{-1}$. Then, the algorithm

$$\delta_{k+1} = \delta_k + \beta\big(\gamma - h(\delta_k)\big), \quad \delta_0 = 0,$$

converges to $\delta : h(\delta) = \gamma$ at a linear rate proportional to the estimation error of $L$. Now, to obtain the feedback $K^\star(\gamma)$,

[6]$K = -\alpha(R + \alpha B^\top P B)^{-1} B^\top P \widehat{A}$ for $P = Q + \alpha \widehat{A}^\top P \widehat{A} - \alpha^2 \widehat{A}^\top P B (R + \alpha B^\top P B)^{-1} B^\top P \widehat{A}$.

[7]$\gamma = 0$ would yield 0 since $\Sigma_x \succ 0$.




(a) For $\gamma \in \Gamma$, the sets $\mathbb{\Delta}_\gamma(\widehat{A} + BK^\star(0))$ with the corresponding worst-case path, including a projection on the cost surface.

(b) Comparison of the uncertainty hedged against for the nominal-, $K^\star(0)$, and robust controller $K^\star(1)$, *i.e.,* $\mathbb{\Delta}_{\gamma=1}$ under both $K$.

Fig. 2: Given the parameters from section V we show the worst-case uncertainties via Proposition III.6 plus how our robust controller anticipates on where the cost increases the sharpest.

given the correct $\delta$, one can solve the Generalized Algebraic Riccati Equation (7) iteratively as proposed in [30].

In the next section we will briefly present the technical proofs and corresponding supporting material.

## VI. Appendix; proofs and supporting material

First we prove Proposition III.4, which is split up in two parts.

*Proof:* [Proof of Proposition III.4 (i)] Let $\widehat{A}_{c\ell} \triangleq \widehat{A} + \widehat{B}K \in \mathbb{R}^{3\times 3}$ and $\Delta_{A_{c\ell}} \in \mathbb{R}^{3\times 3}$ be parametrized by $\alpha \in (0,1)$ and the finite scalars $(a,b,c,d)$ with $d \in (-1,1)$:

$$\widehat{A}_{c\ell} = \frac{1}{\sqrt{\alpha}}\begin{pmatrix} d & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{pmatrix}, \quad \Delta_{A_{c\ell}} = \frac{1}{\sqrt{\alpha}}\begin{pmatrix} 0 & 0 & a \\ b & 0 & c \\ 0 & 0 & 0 \end{pmatrix}.$$

By construction all these $\widehat{A}_{c\ell} + \Delta_{A_{c\ell}}$'s are $\sqrt{\alpha}$-stable. Say we want $\Delta_{A_{c\ell}}$ and $\Delta_{A_{c\ell}}^\top$ to be in some $\mathbb{\Delta}_\gamma(\widehat{A}_{c\ell})$. Then for simplicity assume $K = D = \Sigma_v = \Sigma_0 = I_3$ such that we only need to find a valid $\gamma$. By stability of both $\widehat{A}_{c\ell} + \Delta_{A_{c\ell}}$ and $\widehat{A}_{c\ell} + \Delta_{A_{c\ell}}^\top$, the matrix $\Sigma_x$ exists for all $\alpha \in (0,1)$ such that we can always find a $\gamma \in \mathbb{R}$ being equal to $\max\{\mathrm{Tr}(\Delta_{A_{c\ell}}^\top \Delta_{A_{c\ell}} \Sigma_{x,\Delta_{A_{c\ell}}}), \mathrm{Tr}(\Delta_A \Delta_A^\top \Sigma_{x,\Delta_{A_{c\ell}}^\top})\}$. So $\Delta_{A_{c\ell}}$ and $\Delta_{A_{c\ell}}^\top$ are members of some $\mathbb{\Delta}_\gamma(\widehat{A}_{c\ell})$. Now let $\Delta_X := \theta\Delta_{A_{c\ell}} + (1-\theta)\Delta_{A_{c\ell}}^\top$, $\theta \in [0,1]$. Then for $\theta = 0.5$ and $a = b = c = 4$, $d = 0.5$ we have $\lambda(\widehat{A}_{c\ell} + \Delta_X) = \alpha^{-1/2}\{-1.5, -1.5, 4.5\}$ such that $\Delta_X \notin \mathbb{\Delta}_\gamma(\widehat{A}_{c\ell})$ since $\Sigma_x \notin \mathcal{S}_+^n$. This example can be generalized to higher dimensions. Since here we have $\Delta_{A_{c\ell}} = \Delta_A + \Delta_B$, one can easily see that for example when $B$ is known, the admissible uncertainties in $A$ might live in a non-convex set.

The set (2) has another interesting property indeed

*Proof:* [Proof of Proposition III.4 (ii)] First, using the Kronecker product ($\otimes$) we rewrite the expression for $\mathcal{A}_\gamma(\widehat{A}_{c\ell})$. Let $W := \alpha(1-\alpha)^{-1}\Sigma_v + \Sigma_0 \succ 0$, then the discrete Lyapunov equation can be represented as $\mathrm{vec}(\Sigma_x) = (I_{n^2} - \alpha A_{c\ell} \otimes A_{c\ell})^{-1}\mathrm{vec}(W)$. Secondly, for $\Delta_{A_{c\ell}} \in \mathbb{R}^{d\times n}$ the inner product becomes:

$$\begin{aligned}
&\langle \Delta_A^\top \Delta_A, \Sigma_x \rangle = \\
&= \mathrm{Tr}(\Delta_{A_{c\ell}}^\top \Delta_{A_{c\ell}} \Sigma_x) = \mathrm{Tr}(\Delta_{A_{c\ell}} \Sigma_x \Delta_{A_{c\ell}}^\top) \\
&= \mathrm{vec}^\top(I_d)\mathrm{vec}(\Delta_{A_{c\ell}} \Sigma_x \Delta_{A_{c\ell}}^\top) \\
&= \mathrm{vec}^\top(I_d)(\Delta_{A_{c\ell}} \otimes \Delta_{A_{c\ell}})\mathrm{vec}(\Sigma_x) \\
&= \mathrm{vec}^\top(I_d)(\Delta_{A_{c\ell}} \otimes \Delta_{A_{c\ell}})(I_{n^2} - \alpha A_{c\ell} \otimes A_{c\ell})^{-1}\mathrm{vec}(W).
\end{aligned}$$

Thus the algebraic equation for $\Sigma_x$ *can* be omitted, but note, at this point we have lost the stability constraint $\Sigma_x \succ 0$. For ease of notation let $D = I_n$, define $Z := I_{n^2} - \alpha(\widehat{A}_{c\ell} + \Delta_{A_{c\ell}}) \otimes (\widehat{A}_{c\ell} + \Delta_{A_{c\ell}})$ and the $\mathrm{mat}(\cdot)$ operator by $X = \mathrm{mat}(\mathrm{vec}(X))$. Then for $Y := \mathrm{mat}(Z^{-1}\mathrm{vec}(W))$ the set $\mathbb{\Delta}_\gamma(\widehat{A}_{c\ell}) \subset \mathbb{R}^{n\times n}$ can be written as

$$\left\{ \Delta_{A_{c\ell}} : \begin{array}{l} 0 \leq \mathrm{vec}^\top(I_n)(\Delta_{A_{c\ell}} \otimes \Delta_{A_{c\ell}})Z^{-1}\mathrm{vec}(W) \leq \gamma \\ 0 < \det(Y_i), \quad i = 1, \ldots, n \end{array} \right\} \tag{13}$$

for $\det(Y_i)$ being the $i^{\text{th}}$ principal minor of $Y$. This additional strictly-positive determinant constraint asserts selection of uncertainties leading to $\sqrt{\alpha}$-stable $A_{c\ell}$ by enforcing $Z \succ 0$, see *e.g.* Theorem 7.2.5 in [31]. Differently put, the principal minor constraints re-enforce $\Sigma_x \succ 0$ again. Using Cramer's rule, *i.e.* $Z^{-1} = \mathrm{adj}(Z)/\det(Z)$, it can be observed that (13) is indeed semi-algebraic for $\gamma > 0$, thus a set of polynomial inequalities in the elements of $\Delta_{A_{c\ell}}$ of the form $\mathcal{S}$:

$$\mathcal{S} = \{\Delta_{A_{c\ell}} \in \mathbb{R}^{d\times n} \; : \; 0 \leq p_1(\Delta_A),$$
$$0 \leq \gamma p_2(\Delta_A) - p_1(\Delta_A), \; 0 < p_i(\Delta_A), \quad i = 3, \ldots, 3+n\}.$$

This result is of course closely related to the prominent role played by polynomials in linear control theory. We can add that thereby, our set is a disjoint union of a finite number of connected semi-algebraic sets, which follows directly from the fact that $\mathcal{A}_\gamma$ is semi-algebraic and Theorem 5.19 in [32].

*Proof:* [Proof of Lemma III.6] Consider the problem

$$\mathcal{P}_a(\gamma): \underset{\Delta_{A_{c\ell}} \in \mathbb{\Delta}_\gamma(\widehat{A}_{c\ell})}{\operatorname{argmax}} \mathcal{J}(\widehat{A}_{c\ell} + D\Delta_{A_{c\ell}}, Q_{c\ell}),$$

If $\gamma$ satisfies $h(\delta) = \gamma$ then from the Lemma VI.1 the solution to $\mathcal{P}_a(\gamma)$ can be directly retrieved from the (negated) problem

$$\mathcal{P}_b(\delta): \begin{cases} \underset{\Delta_{A_{c\ell}} \in \mathbb{R}^{d\times n}}{\operatorname{argmin}} & \underset{x_0,v}{\mathbb{E}} \left[ \sum_{k=0}^{\infty} \alpha^k \left( \delta^{-1} w_k^\top w_k - x_k^\top Q_{c\ell} x_k \right) \right] \\ \text{subject to} & x_{k+1} = \widehat{A}_{c\ell} x_k + D w_k + v_k, \\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0, \Sigma_v),\ x_0 \sim \mathcal{P}(0, \Sigma_0), \\ & w_k = \Delta_{A_{c\ell}} x_k. \end{cases} \tag{14}$$

Under the conditions from Proposition III.6 the program $\mathcal{P}_b(\delta)$ can be solved using Dynamic Programming, *e.g.* see chapter 3 from [33], regarding feasibility one can always select $w_k = 0\ \forall k$, moreover $(\delta^{-1} I_d - \alpha D^\top S D) \succ 0$ asserts boundedness of the cost from below. Let the Value function (cost-to-go from state $x$, *i.e.,* without taking the expectation over $x_0$), corresponding to (14), under a policy $\nu := \{w_0, w_1, \dots\}$ be parameterized by $V^\nu(x) = -x^\top S x + q$, $S \in \mathcal{S}^n_+$, $q \in \mathbb{R}$. An expression for the optimal policy and value function follow from the classical Bellman equation

$$V^\nu(x) = \inf_\nu \left\{ c(x,w) + \alpha \mathbb{E}_{x' \sim \mathcal{P}(\cdot|x,\nu(x))} \left[ V^\nu(x') \right] \right\},$$

which yields in the context of (14)

$$\begin{aligned} &-x^\top S x + (1-\alpha) q \\ &= \inf_w \big\{ \delta^{-1} w^\top I_d w - x^\top Q_{c\ell} x \\ &\quad - \alpha \mathbb{E}_v \left[ (\widehat{A}_{c\ell} x + D w + v)^\top S (\widehat{A}_{c\ell} x + D w + v) \right] \big\} \\ &= \inf_w \Bigg\{ \begin{pmatrix} x \\ w \end{pmatrix}^\top \Bigg[ \begin{pmatrix} -Q_{c\ell} & 0 \\ 0 & \delta^{-1} I_d \end{pmatrix} \\ &\quad - \alpha \begin{pmatrix} \widehat{A}_{c\ell}^\top S \widehat{A}_{c\ell} & \widehat{A}_{c\ell}^\top S D \\ D^\top S \widehat{A}_{c\ell} & D^\top S D \end{pmatrix} \Bigg] \begin{pmatrix} x \\ w \end{pmatrix} - \alpha \operatorname{Tr}(S \Sigma_v) \Bigg\} \\ &= x^\top \big( -Q_{c\ell} - \alpha \widehat{A}_{c\ell}^\top S \widehat{A}_{c\ell} \\ &\quad - \alpha^2 \widehat{A}_{c\ell}^\top S D (\delta^{-1} I_d - \alpha D^\top S D)^{-1} D^\top S \widehat{A}_{c\ell} \big) x \\ &\quad - \alpha \operatorname{Tr}(S \Sigma_v), \end{aligned}$$

if $(\delta^{-1} I_d - \alpha D^\top S D) \succ 0$ indeed. Thus, the optimal policy is

$$w_k^\star = \alpha (\delta^{-1} I_d - \alpha D^\top S D)^{-1} D^\top S \widehat{A}_{c\ell} x_k,$$

where

$$\begin{aligned} S =& Q_{c\ell} + \alpha \widehat{A}_{c\ell}^\top S \widehat{A}_{c\ell} \\ &+ \alpha^2 \widehat{A}_{c\ell}^\top S D (\delta^{-1} I_d - \alpha D^\top S D)^{-1} D^\top S \widehat{A}_{c\ell}, \end{aligned}$$

resembles the corresponding Riccati equation. This directly gives the expression for $\Delta^\star_{A_{c\ell}}(\delta)$ and concludes the proof.

*Proof of Theorem III.7:* Now this apparent link between the solution to a robust LQR problem and a dynamic game as set forth in section IV is formalized, which constitutes the main result, Theorem III.7. This is not new, see for example [16], [27], where in the latter[8], the pair $(\gamma, \delta)$ is interpreted via multiplier theory (*cf.* [34], [35]) with respect to a constraint of the form $\sum_{k=0}^\infty \alpha^k w_k^\top w_k \leq \gamma$. We provide a slightly different proof in terms of $(K, L)$ instead of $\big(\{u_k\}_k, \{w_k\}_k\big)$ which eventually allows for numerically finding a solution depending on $\delta$, given $\gamma$ (see Remark V.1).

Recall Definition III.1 and the RLQR problem (1). Let a solution to (3) be denoted by the pair $\big(K^\star(\gamma), \Delta_A^\star(\gamma)\big)$ whereas a solution to (12), if it exists, is $\big(K^\star(\delta), L^\star(\delta)\big)$. Then the next proof allows us to link the solution from the dynamic game (12) to the solution of the robust LQ regulator (3). This proof of Theorem III.7 is split up into a few parts.

*Proof:* [Proof of Theorem III.7 part (i),(ii) and (iii)] Regarding the monotonicity in (iii), first consider the game (12). By Lemma VI.2 the cost can be equivalently written as $f(K,L) - \delta^{-1} g(K,L)$ for $u_k = K x_k$, $w_k = L x_k$, $x_{k+1} = A x_k + B u_k + D w_k + v_k$ and the pair $f(K,L)$, $g(K,L)$ being defined by

$$f(K,L) = \underset{x_0,v}{\mathbb{E}} \left[ \sum_{k=0}^\infty \alpha^k x_k^\top \left( Q + K^\top R K \right) x_k \right], \tag{15}$$

$$g(K,L) = \underset{x_0,v}{\mathbb{E}} \left[ \sum_{k=0}^\infty \alpha^k w_k^\top w_k \right] = \left\langle L^\top L, \Sigma_x \right\rangle, \tag{16}$$

with $\Sigma_x = \underset{x_0,v}{\mathbb{E}} \left[ \sum_{k=0}^\infty \alpha^k x_k x_k^\top \right]$[9]. Then $\sup_L \{ f(K', L) - \delta^{-1} g(K', L) \}$ corresponds to program $\mathcal{P}_2$ from Lemma VI.1 with the map $h$ from (6) and an additional (fixed) parameter $K'$. The map $h(\delta)$ is non-decreasing on some interval $[0, \overline{\delta}) \subset \mathbb{R}_{\geq 0}$, $\overline{\delta} < \infty$. To see why we have this interval, recall that feasibility of the game is defined by a condition of the form $\delta : \delta^{-1} I - P \succ 0$. Indeed, in [11], [16] the parameter $\overline{\delta}$ resembles their "*breakdown*" point $\underline{\theta}$.

Regarding (i)-(ii), by construction of the result for (iii), the programs (3) and (12) are of the form

$$\widetilde{\mathcal{P}}_1(\gamma): \begin{cases} \inf_{K \in \mathbb{R}^{m\times n}} & \sup_{L \in \mathbb{R}^{d\times n}} f(K,L) \\ & \quad \text{s.t.} \quad g(K,L) \leq \gamma, \end{cases}$$

$$\widetilde{\mathcal{P}}_2(\delta): \inf_{K \in \mathbb{R}^{m\times n}} \sup_{L \in \mathbb{R}^{d\times n}} f(K,L) - \delta^{-1} g(K,L),$$

respectively, for $f(K,L)$ and $g(K,L)$ defined by (15) and (16).

These programs $\big(\widetilde{\mathcal{P}}_1(\gamma), \widetilde{\mathcal{P}}_2(\delta)\big)$ correspond to $\mathcal{P}_1(\gamma)$ and $\mathcal{P}_2(\delta)$ from Lemma VI.1 but with an outer minimization step over $K$. Let the corresponding solutions to the inner maximazition problems be denoted by $L_1^\star(\gamma, K)$ and $L_2^\star(\delta, K)$. Then by Lemma VI.1 we have $L_1^\star(\gamma, K) = L_2^\star\big(h^{-1}(\gamma), K\big)$.

[8] Specifically, see sec. 2.4 for an introduction and ch.7 and 8 for a formal discussion.

[9] This step relies on the Bounded Convergence Theorem (*cf.* p.57 [36]) in that implicit in the definition of $h(\delta)$ resides feasibility of the game, thereby boundedness of the two parts of the cost. This justifies the splitting of $\mathbb{E}[\cdot]$, *i.e.,* $\lim_{n\to\infty} \int_{\mathcal{X}} f_n + g_n \mathrm{d}\mu = \int_{\mathcal{X}} \lim_{n\to\infty} f_n \mathrm{d}\mu + \int_{\mathcal{X}} \lim_{n\to\infty} g_n \mathrm{d}\mu$.

Moreover, when $h(\delta) = \gamma$ then $L_1^\star(\gamma, K) = L_2^\star(\delta, K)$ and thereby $g\big(K, L_1^\star(\gamma, K)\big) = g\big(K, L_2^\star(\delta, K)\big)$.

Now let $K^\star(\delta)$ be the solution to the outer minimization of $\widetilde{\mathcal{P}}_2$. To show that this $K^\star(\delta)$ is also optimal for $\widetilde{\mathcal{P}}_1$ assume, like in Lemma VI.1 for the sake of contradiction it is not. For $\widetilde{\mathcal{P}}_1$ we effectively consider $\inf_K \left\{ f\big(K, L_1^\star(\gamma, K)\big)\right\}$ where it is known that $g\big(K, L_1^\star(\gamma, K)\big) \leq \gamma$ holds. However, since $h(\delta) = \gamma$ we can equivalently consider $\inf_K \left\{ f\big(K, L_2^\star(\delta, K)\big)\right\}$. Then to continue the contradictive argument assume there is some $\widetilde{K}$ such that

$$f\big(\widetilde{K}, L_2^\star(\delta, \widetilde{K})\big) < f\Big(K^\star(\delta), L_2^\star\big(\delta, K^\star(\delta)\big)\Big).$$

By construction we have $h(\delta) = \gamma$, and thus $g\big(\widetilde{K}, L_2^\star(\delta, \widetilde{K})\big) = \gamma = g\Big(K^\star(\delta), L_2^\star\big(\delta, K^\star(\delta)\big)\Big)$ such that existence of such a $\widetilde{K}$ contradicts optimality of $K^\star(\delta)$ in $\widetilde{\mathcal{P}}_2$. Therefore, the condition that $h(\delta) = \gamma$ implies that if the pair $\big(K^\star(\delta), L^\star(\delta)\big)$ exists, it is an optimal solution to both (12) and (3).

Thus, when there is a $\delta \geq 0 : h(\delta) = \gamma$, which we have by construction of the Theorem, then the solution to (3) is given by the pair $\big(K^\star(\delta), L^\star(\delta)\big)$, for which the expressions are given by Lemma VI.2. Moreover, the statement of the Theorem can be extended to assert that these matrices exist, as the conditions can be made to be in correspondence with this Lemma VI.2 (feasibility of (12), *e.g.,* $(A, B, C)$ being a minimal realization).

At last we characterize the regularity of the map $h$ in the context of Theorem III.7, which is again very useful with numerical algorithms in mind. This is done in the spirit of the work by Polderman [37], [38].

*Proof:* [Proof of Theorem III.7 (iii) cont.] We will first show that $\overline{P}^+(\delta)$[10] is analytic over $[0, \bar{\delta})$, whereafter the result easily follows via the dependence of $h(\delta)$ on $P(\delta)$. Let $C$ be defined by $Q = C^\top C$. Then define for an arbitrary minimal realization $(A, B, C)$ the matrix valued map $\ell : \mathbb{R}_{\geq 0} \times \mathcal{S}_+^n \to \mathcal{S}_+^n$ by

$$\begin{aligned}\ell(\delta, P) =& P - Q - \alpha A^\top P \cdots \\ & \cdots \left(I_n + \alpha\left(BR^{-1}B^\top - \delta DD^\top\right)P\right)^{-1} A.\end{aligned} \tag{17}$$

This map $\ell$ is $C^\omega$ over some open set $(0, \bar{\delta}) \times V \subset \mathbb{R}_{\geq 0} \times \mathcal{S}_+^n$ since rational functions are analytic on their domain. To continue, we will show that in specific neighbourhoods of $(\tilde{\delta}, \widetilde{P}) \in (0, \bar{\delta}) \times V$, zeroing $\ell$, there exist $C^\omega$ maps $P(\delta)$ such that $\ell(\delta, P(\delta)) = 0$. To that end, define $\Gamma(\Delta_P) \triangleq \ell(\tilde{\delta}, \widetilde{P} + \Delta_P)$ and consider only the linear terms, denoted by $\overset{L}{=}$, in $\Delta_P$:

$$\begin{aligned}\Gamma(\Delta_P) \overset{L}{=}& \Delta_P - \alpha A^\top(\widetilde{P} + \Delta_P)\cdots \\ & \cdots \left(I_n + \alpha\left(BR^{-1}B^\top - \tilde{\delta}DD^\top\right)(\widetilde{P} + \Delta_P)\right)^{-1} A \\ \overset{L}{=}& \Delta_P - \alpha A^\top(\widetilde{P} + \Delta_P)\widetilde{\Lambda}^{-1}\sum_{k=0}^{\infty}(-1)^k \cdots \\ & \cdots \left(\alpha(BR^{-1}B^\top - \tilde{\delta}DD^\top)\Delta_P\widetilde{\Lambda}^{-1}\right)^k A \\ \overset{L}{=}& \Delta_P - \alpha A^\top \cdots \\ & \cdots (I_n - \widetilde{P}\widetilde{\Lambda}^{-1}\alpha(BR^{-1}B^\top - \tilde{\delta}DD^\top)\Delta_P\widetilde{\Lambda}^{-1}A \\ \overset{L}{=}& \Delta_P - \alpha A^\top\widetilde{\Lambda}^{-\top}\Delta_P\widetilde{\Lambda}^{-1}A.\end{aligned}$$

These steps hinge on geometric series for matrices, and a few linear algebraic identities[11]. Now since we know that $\widetilde{\Lambda}^{-1}A$ is $\sqrt{\alpha}$-stable when $\widetilde{P}$ is $\overline{P}^+(\tilde{\delta})$, the map $\Gamma$ must be non-singular (see Lemma 2.3 [37]) for such a point $\big(\tilde{\delta}, \overline{P}^+(\tilde{\delta})\big)$. Therefore, we can apply the Implicit Function Theorem (*cf.* [39]), which asserts (locally) the existence of an unique $C^\omega$ map $P(\delta)$ such that $\ell(\delta, P(\delta)) = 0$ and $P(\delta) = \widetilde{P}$ for all $\delta \in U_\delta \subset \mathbb{R}_{\geq 0}$. Since the pair $(\tilde{\delta}, \widetilde{P})$ was arbitrary, up to being a minimal solution, this holds for any pair $(\delta, \overline{P}^+(\delta))$, making $\overline{P}^+(\delta) \in C^\omega\big((0, \bar{\delta})\big)$. This implies that $L^\star(\delta)$ is $C^\omega$ in $\delta$ and by Theorem E.1.4[12]. from [40], so is $\Sigma_x$, such that indeed the map $h(\delta)$ is analytic over some bounded interval. Finally, to extend $(0, \bar{\delta})$ to $[0, \bar{\delta})$ observe that $\lim_{\delta \downarrow 0} h(\delta) = 0$, which concludes the proof.

The following lemma is the key to bridge the RLQR problem (1) under uncertainty sets from Definition III.1 to a dynamic game theory perspective.

**Lemma VI.1** (Exact constraint relaxation)**:** Let $f, g$ be functions from $\mathcal{X}$ to $\mathbb{R} \cup \{\infty\}$. Given a parameter $\gamma \geq 0$, we define the optimization programs

$$\mathcal{P}_1(\gamma) : \begin{cases} \sup\limits_{x \in \mathcal{X}} & f(x) \\ \text{s.t.} & g(x) \leq \gamma, \end{cases} \qquad \mathcal{P}_2(\gamma) : \sup_{x \in \mathcal{X}} f(x) - \gamma^{-1}g(x),$$

where $x_i^\star(\gamma)$, $i \in \{1, 2\}$, denote an optimizer of the corresponding program. Then, the following holds:

(i) The function $h(\gamma) := g\big(x_2^\star(\gamma)\big)$ is non-decreasing over $\gamma \in \mathbb{R}_{\geq 0}$ when $\mathcal{P}_2(\gamma)$ admits an optimal solution.

(ii) A solution to the program $\mathcal{P}_1(\gamma)$ can be retrieved via $x_1^\star(\gamma) = x_2^\star\big(h^{-1}(\gamma)\big)$, where $h^{-1}$ denotes the inverse function of $h$ defined in (i).[13]

*Proof:* Consider the parameters $\gamma_1 \geq \gamma_2$, and let $x_2^\star(\gamma_1)$ and $x_2^\star(\gamma_2)$ be the optimizers of the program $\mathcal{P}_2$, respectively. In view of the optimality of these solutions, one can readily deduce that

$$\begin{aligned} f\big(x_2^\star(\gamma_1)\big) - \gamma_1^{-1}g\big(x_2^\star(\gamma_1)\big) &\geq f\big(x_2^\star(\gamma_2)\big) - \gamma_1^{-1}g\big(x_2^\star(\gamma_2)\big) \\ f\big(x_2^\star(\gamma_2)\big) - \gamma_2^{-1}g\big(x_2^\star(\gamma_2)\big) &\geq f\big(x_2^\star(\gamma_1)\big) - \gamma_2^{-1}g\big(x_2^\star(\gamma_1)\big). \end{aligned}$$

[10] See Lemma VI.2 for more on this notation.

[11] Most notably: $P(1 + QP)^{-1} = (1 + PQ)^{-1}P$ and $(I + P)^{-1} = I - (I + P)^{-1}P$.

[12] Effectively, by the results from Polderman [38]

[13] In case the inverse function has more than one solution, any selection from the set $h^{-1}(\gamma)$ fulfills the assertion of (ii).

Adding the two sides of the above inequalities yields

$$(\gamma_2^{-1} - \gamma_1^{-1}) g\big(x_2^\star(\gamma_2)\big) \le (\gamma_2^{-1} - \gamma_1^{-1}) g\big(x_2^\star(\gamma_1)\big) \iff g\big(x_2^\star(\gamma_2)\big) \le g\big(x_2^\star(\gamma_1)\big)$$

which concludes the assertion (i).

For (ii), we first argue that any optimal solution to $\mathcal{P}_2(\gamma)$ is an optimal solution to $\mathcal{P}_1\Big(g\big(x_2^\star(\gamma)\big)\Big)$, *i.e.*, using the notation of the optimizers, we have $x_2^\star(\gamma) = x_1^\star\Big(g\big(x_2^\star(\gamma)\big)\Big)$ for any $\gamma \ge 0$. To this end, observe that by the definition the optimizer $x_2^\star(\gamma)$ is a feasible solution to the program $\mathcal{P}_1$ when the parameter $\gamma$ is set to $g\big(x_2^\star(\gamma)\big)$. It then suffices to prove the optimality. For the sake of contradiction, assume that there exists a $\widetilde{x}_1 \in \mathcal{X}$ such that $f(\widetilde{x}_1) > f\big(x_2^\star(\gamma)\big)$ and $g(\widetilde{x}_1) \le g\big(x_2^\star(\gamma)\big)$. Under this assumption, we then have

$$f(\widetilde{x}_1) - \gamma^{-1} g(\widetilde{x}_1) > f\big(x_2^\star(\gamma)\big) - \gamma^{-1} g\big(x_2^\star(\gamma)\big)\,,$$

which contradicts the optimality condition of $x_2^\star(\gamma)$ in the program $\mathcal{P}_2$. Thus, we conclude that $x_2^\star(\gamma) = x_1^\star\Big(g\big(x_2^\star(\gamma)\big)\Big)$. Finally, in the light of the inverse function definition (*i.e.*, $\widetilde{\gamma} = h(\gamma)$ if and only if $\gamma \in h^{-1}(\widetilde{\gamma})$), we arrive at the desired assertion $x_2^\star\big(h^{-1}(\widetilde{\gamma})\big) = x_1^\star(\widetilde{\gamma})$. This concludes the proof of (ii).

This first Lemma summarizes the key results we need regarding the dynamic game
(12).

**Lemma VI.2** (*cf.* chapter 3 from [28] for the undiscounted deterministic case)**:** Given a game (12) for $\alpha \in (0,1)$, let $Q \succeq 0, R \succ 0$, $(\sqrt{\alpha}A, B)$ be stabilizable and $(\sqrt{\alpha}A, C)$ detectable for $Q = C^\top C$. If $\delta \in \mathbb{R}_{\ge 0}$ satisfies $(\delta^{-1} I_d - \alpha D^\top P D) \succ 0$[14], where $P$ is the minimal[15] positive semi-definite solution to the Generalized Algebraic Riccati Equation (GARE):

$$\begin{aligned} P =& Q + \alpha A^\top P \Lambda^{-1} A, \\ \Lambda =& \left(I_n + \alpha\left(BR^{-1}B^\top - \delta DD^\top\right) P\right), \end{aligned} \tag{18}$$

then the optimal[16] strategies are time-invariant, linear in $x_k$ for $K^\star(\delta) \in \mathbb{R}^{m\times n}, L^\star(\delta) \in \mathbb{R}^{d\times n}$ and given by

$$\begin{aligned} \nu_k^\star(x_k) &= \alpha\delta D^\top P \Lambda^{-1} A x_k = L^\star(\delta) x_k, \\ \mu_k^\star(x_k) &= -\alpha R^{-1} B^\top P \Lambda^{-1} A x_k = K^\star(\delta) x_k. \end{aligned}$$

Moreover, under these strategies the closed-loop system $(\Lambda^{-1}A)$ is $\sqrt{\alpha}$-stable and the optimal cost is given by $\mathcal{J}^\star = \langle P, \Sigma_0\rangle + \alpha(1-\alpha)^{-1}\langle P, \Sigma_v\rangle$.

[14] An equivalent condition as promoted by [16] is to check $\text{logdet}(\delta^{-1} I_d - \alpha D^\top P D) > -\infty$

[15] In the terminology of p.81 ch.3 [28], given the feasible iterative scheme $P_{k+1} = Q + A^\top P_k \Lambda_k^{-1} A$, $P_0 = Q$. Then call $\overline{P}^+ := \lim_{k\to\infty} P_k$ the minimal solution to the GARE. This distinction between solutions is important since other solutions might exist, which do not give rise to the desired stability properties.

[16] Not a general saddle-point (see *e.g.* [41])

## BIBLIOGRAPHY


[1] K. Astrom and B. Wittenmark, "On self tuning regulators," *Automatica*, vol. 9, pp. 185–199, 1973.

[2] C.-N. Fiechter, "Pac adaptive control of linear systems," in *Proceedings of the Tenth Annual Conference on Computational Learning Theory*, ser. COLT '97, ACM, 1997, pp. 72–80.

[3] M. C. Campi and P. R. Kumar, "Adaptive linear quadratic gaussian control: The cost-biased approach revisited," *SIAM J. CONTROL OPTIM.*, vol. 36, no. 6, pp. 1890–1907, 1998.

[4] J. C. Doyle, "Guaranteed margins for lqg regulators," *IEEE Transactions on Automatic Control*, vol. AC-23, no. 4, pp. 756–757, 1978.

[5] U. Shaked, "Guaranteed stability margins for the discrete-time linear quadratic optimal regulator," *IEEE Transactions on Automatic Control*, vol. 31, no. 2, pp. 162–165, 1986.

[6] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, ser. Oxford Science Publications. Oxford University Press, 1995.

[7] J.-G. Sun, "Perturbation theory for algebraic riccati equations," *SIAM J. MATRIX ANAL. APPL.*, vol. 19, no. 1, pp. 39–65, 1998.

[8] S. Poljak and J. Rohn, "Checking robust nonsingularity is np-hard," *Mathematics of Control, Signals and Systems*, vol. 6, no. 1, pp. 1–9, 1993.

[9] R. D. Braatz, P. M. Young, J. C. Doyle, and M. Morari, "Computational complexity of $\mu$ calculation," in *1993 American Control Conference*, 1993, pp. 1682–1683.

[10] M. de Oliveira, J. Bernussou, and J. Geromel, "A new discrete-time robust stability condition," *Systems & Control Letters*, vol. 37, no. 4, pp. 261 –265, 1999.

[11] P. Whittle, *Risk-sensitive Optimal Control*. Wiley, 1990.

[12] P. P. Khargonekar, I. R. Petersen, and K. Zhou, "Robust stabilization of uncertain linear systems: Quadratic stabilizability and $h^\infty$ control theory," *IEEE Transactions on Automatic Control*, vol. 35, no. 3, pp. 356–361, 1990.

[13] G. Garcia, J. Bernussou, and D. Arzelier, "Robust stabilization of discrete-time linear systems with norm-bounded time-varying uncertainty," *Systems & Control Letters*, vol. 22, no. 5, pp. 327 –339, 1994.

[14] M. Abeille and A. Lazaric, "Improved regret bounds for thompson sampling in linear quadratic control problems," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80, PMLR, 2018, pp. 1–9.

[15] M. Simchowitz, H. Mania, S. Tu, M. I. Jordan, and B. Recht, "Learning without mixing: Towards a sharp analysis of linear system identification," in *Proceedings of the 31st Conference On Learning Theory*, vol. 75, PMLR, 2018, pp. 439–473.

[16] L. P. Hansen and T. J. Sargent, *Robustness*. Princeton University Press, 2007.

[17] I. R. Petersen, V. Ugrinovskii, and A. Savkin, *Robust Control Design Using $H^\infty$ Methods*. Springer, 2000.

[18] Y. Abbasi-Yadkori and C. Szepesvri, "Regret bounds for the adaptive control of linear quadratic systems," *JMLR: Workshop and Conference Proceedings 19, 24th Annual Conference on Learning Theory*, 2011.

[19] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Foundations of Computational Mathematics*, 2019.

[20] S. Tu, "Sample complexity bounds for the linear quadratic regulator," PhD thesis, EECS Department, University of California, Berkeley, 2019.

[21] A. Cohen, T. Koren, and Y. Mansour, "Learning linear-quadratic regulators efficiently with only $\sqrt{t}$ regret," in *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, 2019, pp. 1300–1309.

[22] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for the linear quadratic regulator," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80, PMLR, 2018, pp. 1467–1476.

[23] B. Gravell, P. Mohajerin Esfahani, and T. Summers, "Learning robust control for LQR systems with multiplicative noise via policy gradient," *arXiv e-prints*, 2019.

[24] W. Jongeneel, T. Summers, and P. Mohajerin Esfahani, "Robust linear quadratic regulator: Exact tractable reformulation (extended version)," 2019. [Online]. Available: {http://www.dcsc.tudelft.nl/~mohajerin/Publications/conference/2019/RLQR_extended.pdf}.

[25] M. Fu and B. Barmish, “Adaptive stabilization of linear systems via switching control,” *IEEE Transactions on Automatic Control*, vol. 31, no. 12, pp. 1097–1103, 1986.

[26] R. Hermann and C. F. Martin, “Applications of algebraic geometry to systems theory - part i,” *IEEE Transasctions on Automatic Control.*, vol. 22, no. 1, pp. 19–25, 1977.

[27] G. Didinsky and T. Basar, “Design of minimax controllers for linear systems with non-zero initial states under specified information structures,” *International Journal of Robust and Nonlinear Control*, vol. 2, no. 1, pp. 1–30, 1992.

[28] T. Basar and P. Bernhard, *$H_\infty$-Optimal Control and Related Minimax Design Problems A Dynamic Game Approach*. Birkhauser, 1995.

[29] B. Barmish, “Stabilization of uncertain systems via linear control,” *IEEE Transactions on Automatic Control*, vol. 28, no. 8, pp. 848–850, 1983.

[30] A. A. Stoorvogel and A. J.T. M. Weeren, “The discrete-time riccati equation related to the $h_\infty$ control problem,” *IEEE Transactions on Automatic Control*, vol. 39, no. 3, pp. 686–691, 1994.

[31] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 1990.

[32] S. Basu, R. Pollack, and M.-F. Roy, *Algorithms in Real Algebraic Geometry*. Springer, 2016.

[33] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, ser. Volume 2, third edition. Athena Scientific, 2007.

[34] D. G. Luenberger, *Optimization by Vector Space Methods*, ser. Wiley Professional Paperback Series. John Wiley & Sons, 1969.

[35] D. P. Bertsekas, *Nonlinear Programming*, ser. Second Edition. Athena Scientific, 1999.

[36] K. B. Athreya and S. N. Lahiri, *Measure Theory and Probability Theory*, ser. Springer Texts In Statistics. Springer, 2006.

[37] J. Polderman, “A note on the structure of two subsets of the parameter space in adaptive control problems,” *Systems & Control Letters*, no. 7, pp. 25–34, 1986.

[38] J. W. Polderman, *Adaptive Control & Identification: Conflict or Conflux*, ser. Phd thesis. 1987.

[39] S. Q. Krantz and H. R. Parks, *The Implicit Function Theorem. History, Theory, and Applications*. Birkhauser, 2003.

[40] J. H. van Schuppen, *Mathematical Control and System Theory of Discrete-Time Stochastic Systems*. 2018.

[41] E. Mageirou, “Values and strategies for infinite time linear quadratic games,” *IEEE Transactions on Automatic Control*, vol. 21, no. 4, pp. 547–550, 1976.