Robust Linear Quadratic Regulator: Exact Tractable Reformulation (extended version)

Wouter Jongeneel, Tyler Summers and Peyman Mohajerin Esfahani

***Abstract*— We consider the problem of synthesizing a control law which minimizes an infinite-horizon discounted quadratic cost subject to a partially unknown noisy linear dynamical system. 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- and input 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. This formulation allows for new structural insights in the benefits of game theoretic robust control.**

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 μ-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.

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]–[22]), 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. In short, we present the proofs from [23] and shed some light on a scenario where our framework appears to be beneficial. To be specific, our contributions are as follows:

(i) We propose a novel family of uncertainty sets for the system- and input matrices (section III-C), and show that the worst-case cost over these sets can be solved efficiently (Proposition III.5).

(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.6). Here, we extend the setting from [23], allowing for uncertainties in B as well.

(iii) At last, using structural properties of our worst-case model (Lemma III.9) we give theoretical- and empirical evidence (see section IV-A) that our robust control law is a natural and computationally attractive alternative

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.

for the nominal control law when the pair (A,B) is identified under ℓ_2-regularized linear least-squares.

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 key difference with [23] is that now B can be uncertain as well. The new uncertainty set and the corresponding main results are presented in section III. In section IV, we illustrate the presented results through several numerical simulations. In section V we conclude the work and highlight potential future work. Section VI contains all technical proofs plus some supporting material.

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, $\|X\|^2_{F,Q}$ is used to denote $\mathrm{Tr}(X^\top QX)$ 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 μ and covariance Σ 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*. Moreover, let f be some linear endomorphism, *e.g.*, Ax, then $W^+(f)$ is the non-trivial exponentially stable eigenspace, *i.e.*, $W^+(f) = \bigoplus_{0<|\lambda|<1} E_\lambda(f)$. Similarly, $W^\infty(f) = \bigoplus_{\lambda=0} E_\lambda(f)$. At last, $\mathsf{GL}^+(n,\mathbb{R})$ denotes the connected component of the general linear group, containing all real-valued $n\times n$-dimensional matrices with strictly positive determinant. Precisely this part of GL preserves orientation. See [24, ch.6] for a formal discussion on orientation. We call a linear automorphism (invertible endomorphism) f *orientation preserving* when the sign of the determinant of the unit cube is invariant under the map f. This preservation is denoted by $\mathrm{Or}(f)=1$, otherwise $\mathrm{Or}(f)=-1$.

II. PRELIMINARIES

In this section the problem at hand is introduced. Here, we elaborate on [23] and consider uncertainties in the pair (A,B).

A. Robust LQR problem

Given the matrices $Q\in\mathcal{S}^n_+, R\in\mathcal{S}^m_{++}$, discount factor $\alpha\in(0,1)$ and $\widehat{A}\in\mathbb{R}^{n\times n}, \widehat{B}\in\mathbb{R}^{n\times m}$, $\Sigma_0,\Sigma_v\in\mathcal{S}^n_{++}$, and $\{v_k\}_{k\in\mathbb{N}}$ being a - usually Gaussian - white noise sequence of independent random variables with zero mean and a time-invariant covariance matrix Σ_v, *i.e.*, $\mathbb{E}[v_i]=0$ and $\mathbb{E}[v_iv_j^\top]=\delta_{ij}\Sigma_v$ 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 some uncertainty set $\mathbb{A}$:

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

Hence, we consider the LQR problem where the pair (A,B) is not precisely known, but known to be described by $A=\widehat{A}+\Delta_A$ and $B=\widehat{B}+\Delta_B$. Here our prior estimate of (A,B) is denoted by $(\widehat{A},\widehat{B})$, whereas $(\Delta_A,\Delta_B)\in\mathbb{A}$ is the pair of uncertainties. This settings natural emerges in identification, where the pair $(\widehat{A},\widehat{B})$ resembles the nominal model and $\mathbb{A}$ is some set where the pair (Δ_A,Δ_B) is known to live in with high probability.

Towards solving (1) we make an assumption on $\{\mu_k\}_k$:

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

Also, 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

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

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.

III. MAIN RESULTS

The main objective of this section is to study implications of a closed-form solution to the RLQR problem (3), as introduced in [23].

A. Introduction of a new uncertainty set

To keep to work self-contained we repeat some definitions and results from [23].

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 $W_{0,v} :=$

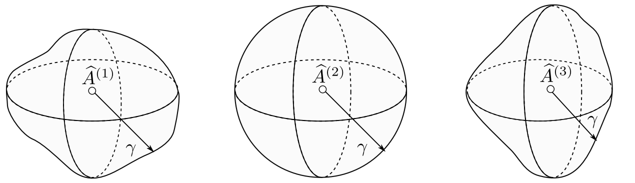


Fig. 1: The set (2) can be interpreted as some ball around $\widehat{A}^{(i)}$. However, for a fixed γ the shape of $\mathcal{A}_\gamma(\widehat{A}^{(i)})$ depends on its center $\widehat{A}^{(i)}$.

$\Sigma_0 + \alpha(1-\alpha)^{-1}\Sigma_v$ and define a set of models around $\widehat{A}$ by the set:

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

For notational convenience, we shall refer to the collection of Δ_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 γ and center $\widehat{A}$. However, in contrast to an Euclidean ball, our set is not translation invariant and depends on the center $\widehat{A}$ (see Figure 1). Moreover, since $W_{0,v} \succ 0$, for Δ_A to be in $\mathbb{\Delta}_\gamma(\widehat{A})$ is the same as being part of the set $\{\Delta_A \in \mathbb{R}^{d\times n} : \|\Delta_A^\top\|^2_{F,\Sigma_x} \le \gamma\}$ for Σ_x as in (2). This further explains why γ 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 $\mathcal{A}_\gamma$.

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

(i) For $n \ge 3$ 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.

At last, when only A is unknown, then using the shorthand notation, the problem (1) over (2) is written as

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

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

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

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

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 $A^\star_{c\ell}(\gamma) := \widehat{A}_{c\ell} + D\Delta^\star_A(\gamma)$.

Proposition III.5 (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 $\delta \in \mathbb{R}_{\ge 0}$, let us assume that $(\delta^{-1}I_d - \alpha D^\top SD) \succ 0$ is satisfied for the (minimal) positive semi-definite solution S to the algebraic equation

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

Then define

$$\Delta^\star_A(\delta) = (\delta^{-1}I_d - \alpha D^\top SD)^{-1}D^\top S\widehat{A}_{c\ell}. \tag{5}$$

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

$$\widetilde{\Sigma}_x = \alpha\big(\widehat{A}_{c\ell} + D\Delta^\star_A(\delta)\big)\widetilde{\Sigma}_x\big(\widehat{A}_{c\ell} + D\Delta^\star_A(\delta)\big)^\top + W_{0,v} \tag{6}$$

which in its turn defines the function

$$\widetilde{h}(\delta) = \left\langle \big(\Delta^\star_A(\delta)\big)^\top\Delta^\star_A(\delta), \widetilde{\Sigma}_x \right\rangle. \tag{7}$$

Then, $\Delta^\star_A(\gamma) = \Delta^\star_A(\delta)$ and $\mathcal{J}^\star = \left\langle \widetilde{\Sigma}_x, Q_{c\ell} \right\rangle$ 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). See [16], [25] for more multiplier interpretations in game theory.

Theorem III.6 (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 $\delta \in \mathbb{R}_{\ge 0}$, assume that the algebraic equation

$$P = Q + \alpha\widehat{A}^\top P\big(I_n + \alpha(BR^{-1}B^\top - \delta DD^\top)P\big)^{-1}\widehat{A} \tag{8}$$

in P admits a minimal[2] 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

$$\Delta^\star_A(\delta) = \alpha\delta D^\top P(\delta)\big(\Lambda(\delta)\big)^{-1}\widehat{A} \tag{9}$$

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 (6) and (7) respectively,

[2]See chapter 3 from [26] for the definition and more information.

which are now functions of K as well, to emphasize the difference, the tildes are dropped, *i.e.,* define:

$$\Sigma_x = \alpha \widehat{A}^\star_{c\ell}(\gamma)\Sigma_x\big(\widehat{A}^\star_{c\ell}(\gamma)\big)^\top + W_{0,v} \tag{10}$$

$$h(\delta) = \Big\langle \big(\Delta^\star_A(\delta)\big)^\top \Delta^\star_A(\delta), \Sigma_x \Big\rangle. \tag{11}$$

Then,

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

$$K^\star(\gamma) = -\alpha R^{-1}B^\top P(\delta)\big(\Lambda(\delta)\big)^{-1}\widehat{A} \tag{12}$$

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 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 $\overline{\delta} < \infty$.

Indeed $\overline{\delta}$ relates to the classical "*breakdown point*" from [11], [16].

Note that we have chosen to interpret (9) as an *additive* uncertainty, but by construction, we could have interpreted the adversarial disturbance as an *multiplicative* uncertainty as well, *e.g.*, $A^\star(\gamma) = \big[I_n + \alpha\delta D^\top P(\delta)\big(\Lambda(\delta)\big)^{-1}\big]\widehat{A} = \Delta \cdot \widehat{A}$. The implications of this observation are discussed in section III-D.

It is also important to stress again that although problem (1) is well-defined for all $\gamma \in \mathbb{R}_{\geq 0}$, Theorem III.6 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 [27, sec. 3-4-2] 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.

C. Uncertainty in the Pair (A, B)

Proposition III.5 and Theorem III.6 are concerned with an uncertainty in the system matrix A. In this section we will show to what extend we can handle uncertainties in B as well, thereby continuing where [23] left off.

1) Classic Approach to Incorporate B: The first approach, as taken in [13], hinges on extending the state space as proposed in [28]. Consider a deterministic dynamical system $x_{k+1} = Ax_k + Bu_k$ and write it in the extended form $x^e_{k+1} = A^e x^e_k + B^e u^e_k$ given by:

$$\begin{pmatrix} x_{k+1} \\ u_{k+1} \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_k \\ u_k \end{pmatrix} + \begin{pmatrix} 0 \\ I_m \end{pmatrix} u^e_k.$$

Now we can appeal to Theorem III.6 with an uncertainty just in A^e, since this block includes uncertainties in both A and B.

To see why this approach is not preferred, let $Q^e = \operatorname{diag}(Q, R)$ and $R^e = \varepsilon I_m \succ 0$ for some $\varepsilon > 0$. Assume that the extended system allows for finding the optimal control gain for $\lim_{\varepsilon\to 0}$ and let the solution be denoted K^e such that $u^e_k := K^e x^e_k$. Back to the original problem, let $u_k := Kx_k$ for some K. Then from $u_{k+1} = K^e_x x_k + K^e_u u_k$ and $u_{k+1} = KAx_k + KBu_k$ we find $K = K_x A^{-1}$ as the solution to the original problem. Although the idea is elegant, this approach has obvious practical obstructions, for example demanding the system matrices to be non-singular.

2) Decompositions of Δ_A: There is another approach to include B within the framework. Let us be given some controller K $\sqrt{\alpha}$-stabilizing $\widehat{A}_{c\ell} \triangleq \widehat{A} + \widehat{B}K$ and $Q_{c\ell} \triangleq Q + K^\top RK$, being the closed-loop cost matrix. Then consider the problem

$$\sup_{A_{c\ell} \in \mathcal{A}_\gamma(\widehat{A}_{c\ell})} \mathcal{J}(A_{c\ell}, Q_{c\ell}), \tag{13}$$

Denote the solution to (4) by $A^\star_{c\ell}(\gamma) \triangleq \widehat{A}_{c\ell} + D\Delta^\star_{A_{c\ell}}(\gamma)$. Now we can directly apply Proposition III.5 and obtain the next Corollary to it.

Corollary III.7 (Decomposition of $\Delta^\star_{A_{c\ell}}(\gamma)$)**:** If (13) is feasible and $\widetilde{h}(\delta) = \gamma$, then worst case uncertainties $\Delta^\star_A(\gamma)$ and $\Delta^\star_B(\gamma)$ are given by

$$\Delta^\star_A(\gamma) = \alpha(\delta^{-1}I_d - \alpha D^\top SD)^{-1}D^\top S\widehat{A}$$
$$\Delta^\star_B(\gamma) = \alpha(\delta^{-1}I_d - \alpha D^\top SD)^{-1}D^\top S\widehat{B}.$$

This follows directly from (5) and $\widehat{A}_{c\ell} = \widehat{A} + \widehat{B}K$ and $\Delta_{A_{c\ell}} \triangleq \Delta_A + \Delta_B K$, but note, this decomposition is not unique.

Under the observation from Corollary III.7 we can write the worst-case closed-loop system as $(I_n + D\Delta^\star)(\widehat{A} + \widehat{B}K)$ for $\Delta^\star := \alpha(\delta^{-1}I_d - \alpha D^\top SD)^{-1}D^\top S$ which is indeed very much in line with equation (36) from [29], although for the infinite-horizon case. The same idea holds for Theorem III.6, consider the problem

$$\inf_{K \in \mathbb{R}^{n\times m}} \sup_{A_{c\ell} \in \mathcal{A}_\gamma(\widehat{A} + \widehat{B}K)} \mathcal{J}(A_{c\ell}, Q + K^\top RK). \tag{14}$$

Assume that (14) is feasible in the sense of Theorem III.6. Then the worst-case model uncertainty, i.e., the maximizing solution to RLQR is $A^\star_{c\ell}(\gamma) = \widehat{A} + \widehat{B}K^\star(\gamma) + D\Delta^\star_{A_{c\ell}}(\delta)$. It turns out that the decomposition of Corollary III.7 carries through:

Lemma III.8 (Decomposition of minimax $\Delta^\star_A(\delta)$)**:** The worst-case uncertainty $\Delta^\star_A(\delta)$ can be decomposed as $\Delta^\star_{A_{c\ell}}(\delta) = \Delta^\star_A(\gamma) + \Delta^\star_B(\gamma)K^\star(\gamma)$ for

$$\begin{aligned} \Delta^\star_A(\gamma) &= \alpha\big(\delta^{-1}I_d - \alpha D^\top P(\delta)D\big)^{-1}D^\top P(\delta)\widehat{A}, \\ \Delta^\star_B(\gamma) &= \alpha\big(\delta^{-1}I_d - \alpha D^\top P(\delta)D\big)^{-1}D^\top P(\delta)\widehat{B}. \end{aligned} \tag{15}$$

Using Lemma III.9.(iii) and VI.5 from below, we clearly see the "*growing*" effect of δ, and by monotonicity in h, of γ. Indeed, γ functions as a radius.

3) An Uncertainty Set for (A, B): Corollary III.7 and Lemma III.8 describe how we can easily decompose closed-loop models and obtain worst-case uncertainties for both the system- and input matrix. The crux is that one can think of $D\Delta_{A_{c\ell}}$ as a perturbation to the nominal system matrix $\widehat{A}$, due to having the same dimension, *or* as sum of perturbations to $\widehat{A}$ and $\widehat{B}$, *e.g.* via $D\Delta_{A_{c\ell}} = D(\Delta_A + \Delta_B K)$. Of course, one could take $\Delta_A \leftarrow \Delta_A + (1-\theta)\Delta_B K$, $\Delta_B \leftarrow \theta\Delta_B$, for any $\theta \in [0,1]$. This interpretation is taken in Example III.5 from [23] for $\theta = 0$, effectively making $\Delta_{A_{c\ell}} \triangleq \Delta_A$. In a special case we also consider some uncertainty *only* in B. If $\exists\ \Delta_B \neq 0 : \Delta_B K^\star = L^\star$ we can define an uncertainty set similar to (2) since the worst-case closed-loop dynamics become $A + (\widehat{B} + D\Delta_B)K$. For example, let $D = \widehat{B}$, then it follows directly from the expressions for $K^\star$ and $L^\star$ (see Lemma VI.2) that $\Delta_B^\star = -\delta R$. Note however that this construction is usually not possible since commonly $m < n$, while $D = I_n$ and L is not rank-deficient.

A final remark on this decomposition is that since we parametrize a dn-dimensional object with $d(n+m)$ parameters, we lose compactness. Moreover, instead of decomposing the solution of Theorem III.6, we could also introduce an uncertainty set for the pair (A,B) directly. Let this set $\mathcal{U}_\gamma \subset \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$ be defined as

$$\mathcal{U}_\gamma\big((\widehat{A},\widehat{B});K\big) = \Big\{(A,B)\ :\ A + BK \in \mathcal{A}_\gamma(\widehat{A} + \widehat{B}K)\Big\}.$$

Then *a* solution to

$$\inf_{K\in\mathbb{R}^{m\times n}} \sup_{(A,B)\in\mathcal{U}_\gamma\left((\widehat{A},\widehat{B});K\right)} \mathcal{J}(A+BK, Q+K^\top RK) \quad (16)$$

is given by (12) and (15). Of course, this description is rather implicit, but it generalizes all the (arbitrary) decompositions from above.

D. Qualitative Properties of the Worst-Case Model

As indicated before, the type of uncertainty set we consider is difficult to *quantify* due to the dependence on $K^\star(\gamma)$[3]. We can however observe several *qualitative* features.

Lemma III.9 (Qualitative features of extremizers in (3), implications of Theorem III.6)**:** For simplicity assume $D = I_n$, then

(i) The worst-case closed-loop system can be written as $\Lambda^{-1}\widehat{A}$ for some $\Lambda^{-1} \in \mathsf{GL}^+(n,\mathbb{R})$, such that the kernel of $\widehat{A}$ is preserved under optimal robust feedback and worst-case uncertainty. Moreover, when $\Sigma_0 \succ 0$ we must have $\mathbb{\Delta}_\gamma\big(\widehat{A} + BK^\star(\gamma)\big) \subseteq \{\Delta_A \in \mathbb{R}^{n\times n} : \mathrm{Ker}(\widehat{A}) \subseteq W^+(\sqrt{\alpha}(\widehat{A} + \Delta_A))\}$ (see Example III.10 below).

(ii) Consider only uncertainty in A, then the automorphic part of the nominal and worst-case drift have the same orientation, *i.e.* $\mathrm{Or}(\widehat{A}x|_a) = \mathrm{Or}\big((\widehat{A} + \Delta_A^\star(\gamma))x|_a\big)$. Moreover, there is a symmetric positive-definite matrix T such that $T\widehat{A} = (\widehat{A} + \Delta_A^\star(\gamma)) = A^\star(\gamma)$, which is stronger than the required $T \in \mathsf{GL}^+(n,\mathbb{R})$ to preserve orientation.

(iii) For $A^\star(\gamma) = \widehat{A} + \Delta_A^\star(\gamma)$, we have $\|A^\star(\gamma)\|_F > \|\widehat{A}\|_F$ almost surely. Moreover, using decomposition (15) we additionally have, *a.s.*, $\|B^\star(\gamma)\|_F > \|\widehat{B}\|_F$.

Orientation is just one part of topological equivalence (*cf.* [30]), but to put it in simple words, item (ii) tells us that an adversarial player does not reveal itself that easily[4]. Moreover, it means that without loss of generality we can optimize over some subset of $\mathcal{A}_\gamma$, preserving the orientation of $\widehat{A}$. What is more, $\mathsf{GL}^{(+)}(n,\mathbb{R})$ is an *invariant* set[5] under $\widehat{A} \mapsto \widehat{A} + \Delta_A^\star(\gamma) =: A^\star(\gamma)$. In addition, item (i) and (ii) imply that $W^\infty(\widehat{A}x)$ is invariant under the worst-case perturbation, with our without feedback. Recall that Least-Squares identification leads (under for example ergodic or episodic assumptions) to ellipsoidal (sublevel) sets of estimates. Then, these observations, together with item (iii), introduce new challenges for *unbiased* identification algorithms, one of them is explained[6] in Figure 2, whereas section IV-A highlights *benefits* in the context of *biased* identification.

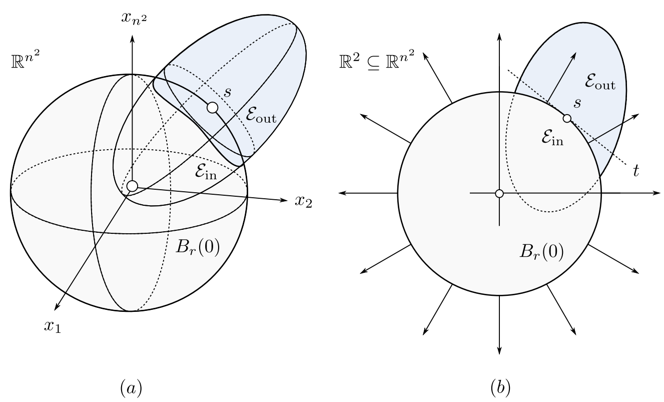


Fig. 2: Let s represent the real system matrix ($\mathrm{vec}(A)$). Using unbiased least-squares we can form an ellipsoid around s given by $\mathcal{E} := \mathcal{E}_{\mathrm{in}} \cup \mathcal{E}_{\mathrm{out}}$, containing estimates of s, denoted $\widehat{s}$. Then since $\|A^\star\|_F > \|\widehat{A}\|_F$ *a.s.*, an estimate $\widehat{s} \in \mathcal{E}_{\min}$ might lead to a worst-case model close to s, while the worst-case model related to a $\widehat{s} \in \mathcal{E}_{\max}$ is even further away from s than the initial $\widehat{s}$. Think of the vectors in (b). The critical observation is however that $\mathrm{Vol}(\mathcal{E}_{\mathrm{out}}) > \mathrm{Vol}(\mathcal{E}_{\mathrm{in}})$ such that the push in the wrong direction is likely to dominate, hence, leading to bad performance. This observation will be further highlighted in section IV-A.

Also, one interpretation for why we have $T \in \mathcal{S}^n_{++}$ is that such a matrix generalizes positive scaling, which is the cheapest method towards destabilization for an adversary.

Next, inspired by [31], we provide an example illustrating the influence of $\widehat{A}$, and its kernel, on optimal LQ regulators.

Example III.10 (Kernel of LQ regulators)**:** Consider the matrices

$$\widehat{A} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad A^\star \begin{pmatrix} 1.1 & 0.1 \\ 0 & 1.2 \end{pmatrix}, \quad B = I_2.$$

[3] See for example (3), we maximize over $\mathcal{A}_\gamma(\widehat{A} + BK)$.

[4] See [27] for further topological implications of Lemma III.9

[5] Hence, setting $\widehat{A} \leftarrow A^\star(\gamma)$ implies that "*worst-worst-case*" models are again members of $\mathsf{GL}^{(+)}(n,\mathbb{R})$. This observation is outside the scope of this work and more interesting in a N-player, $N > 2$, game theoretic framework.

[6] To see how $\mathrm{Vol}(\mathcal{E}_{\mathrm{out}})/\mathrm{Vol}(\mathcal{E}_{\mathrm{in}})$ grows with dimension n, consider two Euclidean balls: $B_r(0)$ and $B_x(0)$ with $x < r/2$. Then from standard volume formulas for n-balls it follows that $\mathrm{Vol}\big(B_r(0)\big)/\mathrm{Vol}\big(B_x(0)\big) > 2^n$ such that $\mathrm{Vol}(\mathcal{E}_{\mathrm{out}})/\mathrm{Vol}(\mathcal{E}_{\mathrm{in}}) \gtrsim C^n$, $C \in (1,2]$.

Then $c(1\ 1)^\top \in \operatorname{Ker}(\widehat{A})\ \forall\ c \in \mathbb{R}$. It is known that any optimal LQ regulator is of the form $K(A,B) = X(A,B)A$. for some non-zero matrix X. Let us design a stabilizing LQR controller for $(\widehat{A}, B)$ and observe that for any non-zero $c \in \mathbb{R}$:

$$\lim_{k\to\infty} \left(A^\star + BX(\widehat{A},B)\widehat{A}\right)^k c \begin{pmatrix}1\\1\end{pmatrix} = \infty.$$

while a simple controller of the form $K = -0.5I_2$ would have done the trick for both $(\widehat{A}, B)$ and $(A^\star, B)$. The key observation is of course that $E_{\lambda=1.2}(A^\star) = \operatorname{Ker}(\widehat{A})$, so that the control gain cannot counteract the growth of the state. This example shows that the usual linear optimal control methods stabilize a very particular subset of systems heavily relient on $\widehat{A}$.

Example III.11 (Vector-Field Interpretation)**:** In practice the matrix $\widehat{A}$ might vary based on incoming data. The aim of this example is to make Figure 2 more concrete and show how the worst-case system matrix uncertainty depends on its center $\widehat{A}$.

Consider for $(x,y) \in [-5,5]^2$ the pair $(\widehat{A}, B)$ and the structural matrix D defined as

$$\widehat{A}(x,y) = \begin{pmatrix}x & y\\0 & 0\end{pmatrix}, \quad B = \begin{pmatrix}0\\1\end{pmatrix}, \quad D = \begin{pmatrix}1\\0\end{pmatrix}. \tag{17}$$

Again, also define the covariance matrices $\Sigma_v = 0.1I_2$, $\Sigma_0 = I_2$, the cost matrices $Q = I_2$, $R = 1$, and the discount factor $\alpha = 0.95$. Now, we compute $\Delta^\star_A(\delta = 10^{-3})$ (9) for each grid-point (x,y) and show the emanating vector (from the first row of $\widehat{A}$ towards the first row of $A^\star(\delta)$). This is done in Figure 3a, where it should be remarked that the arrows solely visualize direction, not tangent vectors of some flow. Around $y = 0$, we lose control, hence no arrows are drawn. More interestingly, see that the vector field is reminiscent of $\dot{z} = z$, $(x,y) =: z \in \mathbb{R}^2$, always pointed away from 0. This follows readily from Lemma III.9 since in this particular case, although $D \neq I_n$, both x and y preserve their sign under being mapped to the worst-case model since

$$A^\star(\delta) \simeq \left[I_2 + \begin{pmatrix}c & 0\\0 & 0\end{pmatrix} Y\right] \begin{pmatrix}x & y\\0 & 0\end{pmatrix}, \quad c \in \mathbb{R}_{>0}, Y \in \mathcal{S}^n_+,$$

and the diagonal elements of a symmetric positive semi-definite matrix are non-negative themselves. This has of course practical implications. For example, in Figure 3b we show 1000 Least-Squares estimates[7] for $A = \widehat{A}(z^\star)$, $z^\star := (x,y) = (1.5, 0.5)$, with the main observation being that indeed the estimates form an ellipsoidal set around this point. Locally, the vector-field is clearly pointing in one direction, which means that if your estimate is for example in the shaded half-space, then the robust control scheme is likely to be ineffective (compare to Figure 2).

In Figure 3c we show that there are indeed a few cases where a robust controller $K^\star(\gamma)$ – for particular γ – improves performance compared to $K^\star(0)$, imagine being in the left halfspace of Figure 3b, moving towards $(1.5, 0.5)$. Nevertheless, on average the performance deteriorates. As will be shown in section IV-A, there are systems and identification settings for which the Least-Squares estimates in combination with our framework *do* improve upon nominal performance.

[7] We used the same procedure as in section IV-A, but with $\lambda = 0$, $N = 10$ and B being known.

IV. Numerical Examples

In this section we highlight via several numerical examples where our framework might be of use. The appendix (section VI-B.2) contains a brief discussion on how to actually carry out the computations involved.

A. Data-Driven Example

Intuitively, it is expected that the robustness coming from a game theoretic approach is useful when one is pessimistic about an estimated model. In other words, the real system should be *worse* in some sense, to be precise, with respect to the cost. In physical systems this occurs for example when inertia is estimated too optimistically, say, when controlling a robotic arm using a model with overestimated inertia. In an abstract setting one can think about marginally stable and sparse models. An estimation scheme might fit stable or dense systems, giving the impression that the controller can relax or has a lot of knobs at its disposal, while in fact, it does not.

To put our framework to the test, we consider almost the same $(n = 3)$-dimensional model as in [32, sec. 4] with $\Sigma_0 = I_3$, $\Sigma_v = 0.1^2 I_3$, $\alpha = 0.95$, $A = \operatorname{tridiag}(0.01, 1.01, 1.01)$, $B = I_3$, $Q = I_3$, $R = I_3$, and give some empirical evidence that our framework can handle these kind of situations. The general setting is as follows: we will do Z experiments, for each experiment z, we let the controlled system run for N steps where we have set[8] $u_k^{(z)} \overset{i.i.d.}{\sim} \mathcal{N}\big(K^\star(0), 0.1^2 I_3\big)$. The resulting data $\{x_k^{(z)}, u_k^{(z)}\}_{k=0}^N$ is the input to a (possibly regularized) Least-Squares problem

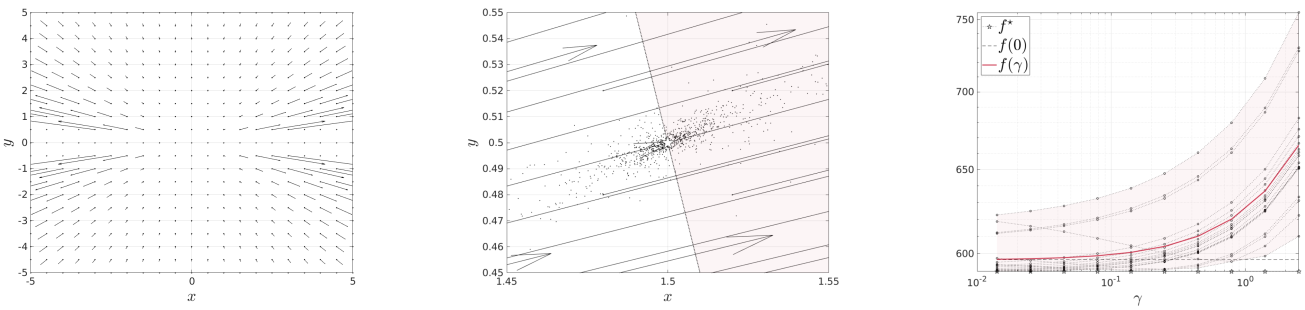
$$(\widehat{A}^{(z)}, \widehat{B}^{(z)}) := \operatorname*{argmin}_{\substack{A\in\mathbb{R}^{n\times n}\\B\in\mathbb{R}^{n\times m}}} \sum_{k=0}^{N-1} \|x_{k+1}^{(z)} - Ax_k^{(z)} - Bu_k^{(z)}\|_2^2 + \lambda\|A\ B\|_F^2$$

which yields a nominal model used for (robust) controller design. Since we have no further structural information, $D = I_3$.

1) Robustness "Sweet-Spot": First, the hope is that if we vary $\gamma \in [0, \overline{\gamma})$, then at some "*radius*", say $\widetilde{\gamma}$, we start including the real system in our uncertainty set, *i.e.,* $\big(A + BK^{(z)\star}(\widetilde{\gamma})\big) \in \mathcal{A}_{\widetilde{\gamma}}\big(\widehat{A}^{(z)} + \widehat{B}^{(z)}K^{(z)\star}(\widetilde{\gamma})\big)$ and tame the real cost, while surpassing performance induced by $K^\star(0)$.

Here we let $Z = 200$, $N = 25$ and $\lambda = 10^{-3}$. It is shown in Figure 4a that we observe precisely this behaviour around $\gamma = 0.08$. When we however increase γ far beyond 10^{-1}, the robust scheme becomes too conservative. We took in total 11 $\gamma \in [0, 2.5]$ and observed that for each

[8] The noise is added to force the input to become persistently exciting.



(a) (b) (c)

Fig. 3: (a) Vector-field corresponding to Example III.11. (b) Zoomed-in version of Figure 3a, together with 1000 Least-Squares estimates of $\widehat{A}(1.5, 0.5)$. (c) Take 20 of the 1000 experiments and let $f^\star$ be the best achievable cost, let $f(0)$ be the empirical mean of the induced cost under $K^\star(0)$ (not a function of γ, merely a reference line) and $f(\gamma)$ the empirical mean of the induced cost under $K^\star(\gamma)$. Moreover, show the 20 individual cost trajectories.

value of γ there is 1 experiment where $K^\star(\gamma)$ fails to $\sqrt{\alpha}$-stabilize the real system. For $\gamma = 2.5$, this value is increased to 3 experiments, the controller became overly pessimistic. Removing the regularization does not change the result structurally, it merely makes the dent (even) less pronounced.

This simple example highlights the potential of our method. Although it must be mentioned that this behaviour is not generic, usually, the robust framework is a lot more conservative. For example, when we reconsider the setup and make B *known*, then the "*sweet-spot*" disappears (see Figure 4b).

Can this behaviour be explained? To that end we recall that in this section we have $D = I_n$ and that Lemma III.9.(iii) pointed out that the worst-case model must be further away from 0 (in Frobenius-norm) than the nominal model. Also recall that from section III-C we know that we can interpret the worst-case system in many was, *e.g.,* as $A^\star = \widehat{A} + \Delta_A^\star$, but we can also think of $\Delta_A^\star$ as $\Delta^\star_{A_{c\ell}} = \Delta_A^\star + \Delta_B^\star K^\star$. Towards understanding the cause of the difference between Figure 4a and Figure 4b let $g_A := Z^{-1}\sum_{z=1}^{Z}\|A\|_F - \|\widehat{A}^{(z)}\|_F$ and $g_{A,B}(\gamma) := Z^{-1}\sum_{z=1}^{Z}\|A + \big(B - \widehat{B}^{(z)}\big)K^{(z)\star}(\gamma)\|_F - \|\widehat{A}^{(z)}\|_F$. Then, it can be shown that for our problem $g_{A,B}(\gamma) > 0$ for all considered γ while $g_A < 0$. Hence, it appears that for our framework to perform well, we need to hope for $\|A\|_F - \|\widehat{A}^{(z)}\|_F \geq 0$, which is precisely in line with the observation made in Figure 2.

Of course, there is a heuristic to enforce $\|A\|_F \geq \|\widehat{A}\|_F$: sufficiently increasing the ℓ_2-*regularization* parameter $\lambda \in \mathbb{R}_{\geq 0}$ to introduce a – for us – favourable bias. However how to select λ? Too small is useless and too big is as if we solve a completely different problem.

The result of increasing $\lambda = 10^{-3}$ to $\lambda = 10^{-1}$ is shown in Figure 4c, and indeed, for a sufficient increase in λ, our framework can still outperform the nominal controller, even when B is known. A remark should be made, introducing (more) regularization does introduce an offset and indeed a higher average nominal cost (and in some examples thereby a higher probability to fail). Nevertheless, it is frequently used in practice to provide some numerical stability (when the normal equations are ill-defined) such that demanding $\lambda > 0$ is far from unrealistic.

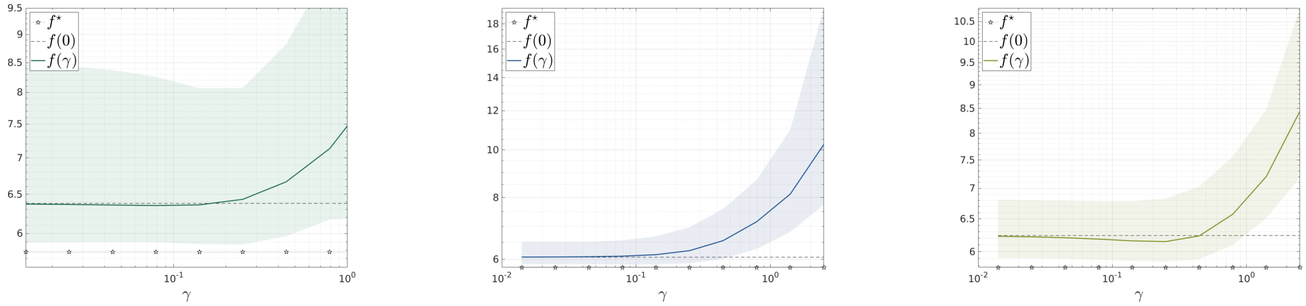
These are marginal improvements, yet based on heuristics, next we investigate the full potential using an optimal selection method.

2) Optimal Selection of γ: Finally, to upper-bound possible performance, we select $\gamma \in \Gamma$ such that $K^{(z)\star}(\gamma)$ achieves the smallest cost on the real system and compare that again to the nominal scenario. Such a γ will be denoted by $\gamma^{(z)\star}$. To start, we let B be known, $\lambda = 0$ and take $N \in \{25, 35, 60, 95, 155\}$ since that is where we expect potential improvement (selecting a smaller N without regularization results in frequent failures and hence a meaningless comparison). The results are shown in Figures 5a,5d and are in line with all simulations before. In fact, when B is known and $\lambda = 0$, then the optimal selection method outperforms the nominal controller just slightly, for $N = 20$, the improvement is exactly 0.15% ($\mathcal{J} = 0.99985 \cdot \mathcal{J}_0$), which decreases along N. Moreover, again in line with Figure 2, $\gamma^{(z)\star} = 0$ is selected for more than 55% of the cases.

However, as before, we can consider some regularization. When we let $\lambda = 10^{-1}$ then the optimal selector can achieve up to 12% cost improvement with respect to the nominal control law, see Figures 5b,5e. Indeed, here we select $\gamma^{(z)\star} = 0$ less than 7% of the time. Note, we zoomed in on a smaller range of N. Also, it is important to recall that we improve with respect to $K^\star(0)$ based on $\widehat{A}$ via regularized Least-Squares, we do not necessarily improve upon $K^\star(0)$ based on non-regularized Least-Squares.

Similarly, we can make B unknown again. These simulation results are shown in Figures 5c,5f. Here, the improvement is at most 2.7% plus we select $\gamma^{(z)\star} = 0$ for more than 55% of the time. Just like before, it is not completely understood why, and when, an uncertain B seems to help, the picture is structurally different than Figure 5b.

Overall, we see that using an unbiased estimator like Least-Squares does not greatly benefit from a clearly biased



(a) Unknown (A, B), $\lambda = 10^{-3}$.

(b) Unknown A, $\lambda = 10^{-3}$.

(c) Unknown A, $\lambda = 10^{-1}$.

Fig. 4: For the Least-Squares procedure from section IV-A.1, discard the best- and worst 5% of the data. Let $f^\star$ be the best achievable cost, let $f(0)$ be the empirical mean (over z) of the induced cost under $K^\star(0)$ (not a function of γ, merely a reference line) and $f(\gamma)$ the empirical mean (over z) of the induced cost under $K^{(z)\star}(\gamma)$. The shaded area is the hull of the remaining 90% of data.

scheme like proposed in this work. Only under sufficient regularization we see some significant improvements.

V. Conclusion and Future Work

"*Systems Identified Under ℓ_2-Regularization Benefit from Game Theoretic Controllers.*" Introducing ℓ_2-regularization into the linear Least-Squares System Identification procedure can have favourable numerical and statistical implications. Especially in the small data-regime is the introduction of $\lambda \in \mathbb{R}_{>0}$ preferred. However, once we use $\lambda > 0$, then the estimates for (A, B) are biased, such that the nominal $K^\star(0)$ is by no means the most natural controller selection anymore. What should we do? By construction we have $\|\widehat{A}|_{\lambda=0} \quad \widehat{B}|_{\lambda=0}\|_F \geq \|\widehat{A}|_{\lambda>0} \quad \widehat{B}|_{\lambda>0}\|_F$. Thus, we would like to select some control law which anticipates on this statistical under-estimation of the Frobenius-norm. Using Lemma III.9.(iii), we see that our robust control law $K^\star(\gamma)|_{\gamma\in(0,\overline{\gamma})}$ is fit for the job since it anticipates on a model being bigger in Frobenius-norm. This concept is summarized in Figure 6 (see Figure 7 for a remark on the *direction* of the arrows).

And indeed, in Figure 5b we observed that regularization helps in the small data regime, in general, regardless of a robust controller. However, the figure also shows that $K^\star(\gamma^\star)$ outperforms the nominal controller *on average* in the small data-regime, which is the most interesting regime[9].

Hence, when the pair (A, B) is identified using ℓ_2-regularized linear Least-Squares, which *is* common practice (see [34] for a wind turbine identification example), then a game theoretic control law $K^\star(\gamma)$ has favourable properties over the nominal $K^\star(0)$ and due to its computational attractive formulation, provides a realistic alternative.

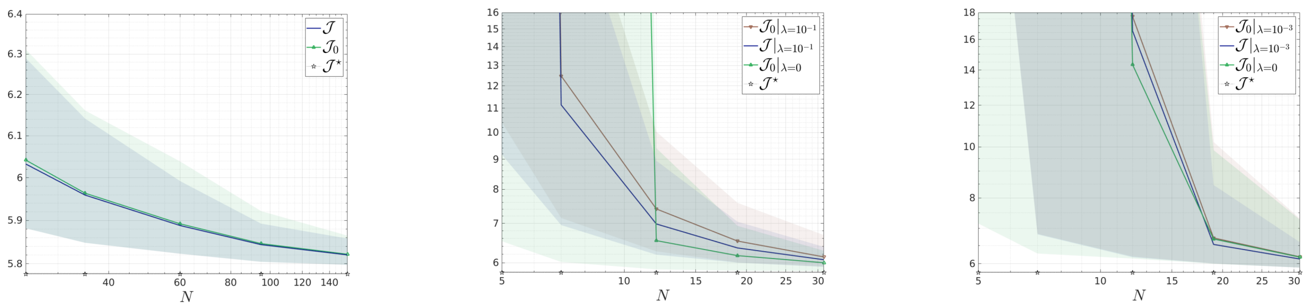
[9] In [33] it is shown that for sufficiently small spectral errors in (A, B) (hence, not the small data-regime), say $\|A - \widehat{A}\| \leq \varepsilon$, the nominal controller is a good choice since the error between the induced cost under the nominal- and best controller scale as $\mathcal{O}(\varepsilon^2)$ (while their robust law scales as $\mathcal{O}(\varepsilon)$). Of course, we saw this performance of $K^\star(0)$ throughout section IV-A.

A. Future Directions

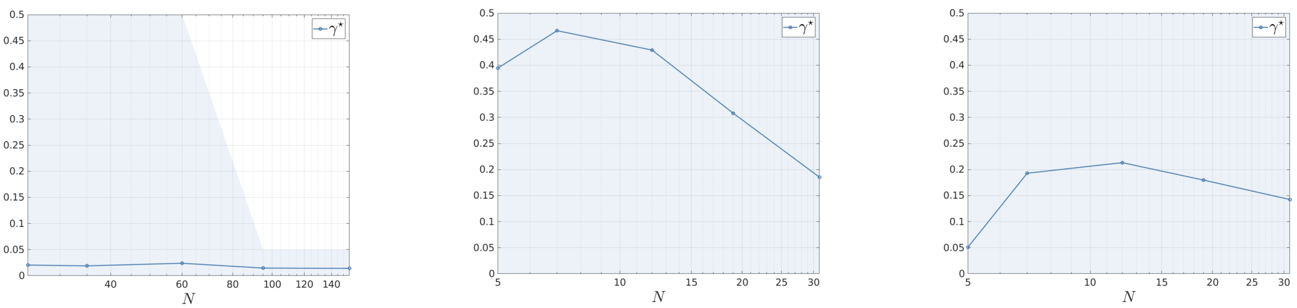
Even in our simple setting there remain many open problems and interesting future research directions. Most notably, can our set be introduced and studied in a full *end-to-end* framework (*cf.* [20])? In other words, can we gain further insights from adaptive schemes for γ? For example, find a map from $\lambda := \lambda'/\sqrt{N}$ to γ and robustly control a system identified under ℓ_2-regularized least-squares. In correspondence with contemporary measure concentration results, we can provide expressions for inscribed norm-balls [27, ch.3], but note again that they are usually inherently small. In line with previous remarks, it might however be more beneficial to first look into the identification algorithms, obtain a better understanding of regularization in our context or look beyond Least-Squares in the first place. Can the observations from section IV-A, especially regarding an uncertain B matrix (see Figure 5c), be further formalized? Recall however that our results are not just empirical, it is especially interesting to note that we sample from a curve (path $p(\gamma)$) in $\mathbb{R}^{m\times n}$, not some ball around $K^\star(0)$ (see Figure 7). Hence, the fact that we improve, on average, cannot be statistical luck. Yet, a map $\psi : (\lambda, N, \Sigma, Q, R, \alpha) \mapsto \gamma$ is missing and would provide significant insights. Additionally, we need a better understanding of the *direction* of the worst-case path; see Figure 7 (b) for an explanation. This strict-halfspace interpretation does not change the previous intuition, but improving upon it lead to a better understanding of the framework.

What is more, can the class of systems giving rise to $\mathcal{A}_\gamma\big(\widehat{A} + \widehat{B}K^\star(\gamma)\big)$ be further formalized as a function of γ (see [27, sec 3.4] for additional examples)? Even more so, it is not investigated, but it is postulated that extensions to the continuous-time, partial-information and distributional regime will bring about new insights. At last, can our approach be of use in other fields relying on (dynamic) game theory, like Reinforcement Learning and Generative Adversarial Networks?

The authors believe further investigations are worthwhile,



(a) Unknown A, $\lambda = 0$. (b) Unknown A, $\lambda = 10^{-1}$. (c) Unknown (A, B), $\lambda = 10^{-3}$.



(d) Unknown A, $\lambda = 0$. (e) Unknown A, $\lambda = 10^{-1}$. (f) Unknown (A, B), $\lambda = 10^{-3}$.

Fig. 5: Select γ optimally (section IV-A.2), discard the top- and bottom 10% of the cost data for $K^\star(\gamma^\star)$ ($\mathcal{J}$) and show, $\gamma^\star$, the best achievable cost ($\mathcal{J}^\star$) plus the cost for $K^\star(0)$ ($\mathcal{J}_0$). All thick lines represent the empirical mean over z. The shaded area is the hull of the remaining 80% of data.

improving our understanding of how to efficiently link identification- and control algorithms towards safe data-driven control.

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

VI. Appendix: proofs and supporting material

A. Proofs

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 $\mathbb{A}_\gamma(\widehat{A}_{c\ell})$. Then for simplicity assume $K = D = \Sigma_v = \Sigma_0 = I_3$ such that we only need to find a valid γ. By stability of both $\widehat{A}_{c\ell} + \Delta_{A_{c\ell}}$ and $\widehat{A}_{c\ell} + \Delta_{A_{c\ell}}^\top$, the matrix Σ_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 $\mathbb{A}_\gamma(\widehat{A}_{c\ell})$. 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 $\Delta_X \notin \mathbb{A}_\gamma(\widehat{A}_{c\ell})$ 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 Σ_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 $X = \mathrm{mat}\big(\mathrm{vec}(X)\big)$. Then for $Y := \mathrm{mat}\big(Z^{-1}\mathrm{vec}(W)\big)$ the set

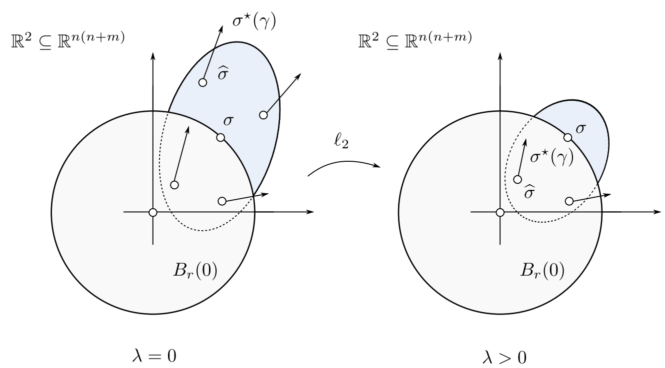


Fig. 6: Let $\sigma := \mathrm{vec}(A\ B)$ be unknown. Using Least-Squares ($\lambda = 0$) obtain an ellipsoidal set around this point. From Figure 2 and Lemma III.9.(iii) we know that the worst-case models $\sigma^\star(\gamma)$, growing from some estimate $\widehat{\sigma}$, move away from 0. Combining this with $\mathrm{Vol}(\mathcal{E}_{\mathrm{in}}) < \mathrm{Vol}(\mathcal{E}_{\mathrm{out}})$ implies that, on average, $\sigma^\star(\gamma)$ is not sufficiently close to σ for the performance to improve upon the nominal control law. However, after introducing ℓ_2-regularization ($\lambda > 0$), the confidence ellipsoid shifts towards 0, plus it becomes more isotropic, hence significantly increasing the probability that $\sigma^\star(\gamma)$ is sufficiently close to σ for some appropriate choice of $\gamma \in (0, \overline{\gamma})$.

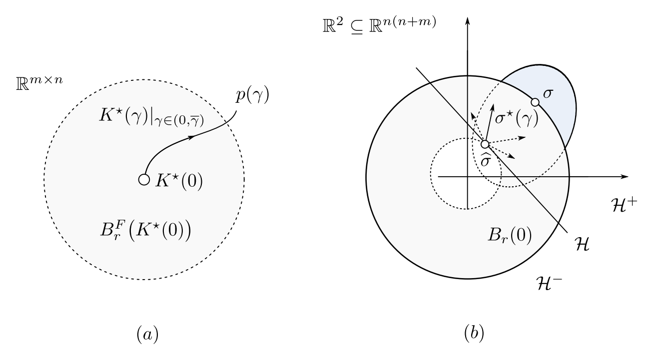


Fig. 7: (a) Throughout the selection methods in section IV-A we select $K^\star(\gamma)$ from a grid on $p(\gamma)$, not some ball $B_r^F\big(K^\star(0)\big)$. Still, we can outperform $K^\star(0)$, which is in favour of the theoretically justified intuition skected out in Figure 6. (b) Copying Figure 6 ($\lambda > 0$) into Figure 7 (b), we drew *a potential* $\widehat{\sigma}^\star(\gamma)$, but all we know from Lemma III.9 is that this vector could have been any of the other dashed arrows pointing in $\mathcal{H}^+ := \mathbb{R}^2 \setminus \{\mathcal{H} \cup \mathcal{H}^-\}$.

$\mathbb{A}_\gamma(\widehat{A}_{\mathrm{c}\ell}) \subset \mathbb{R}^{n\times n}$ can be written as

$$\left\{ \Delta_{A_{\mathrm{c}\ell}} : \begin{matrix} 0 \leq \mathrm{vec}^\top(I_n)(\Delta_{A_{\mathrm{c}\ell}} \otimes \Delta_{A_{\mathrm{c}\ell}})Z^{-1}\mathrm{vec}(W) \leq \gamma \\ 0 < \det(Y_i), \quad i = 1, \dots, n \end{matrix} \right\} \tag{18}$$

for $\det(Y_i)$ being the i^{th} principal minor of Y. This additional strictly-positive determinant constraint asserts selection of uncertainties leading to $\sqrt{\alpha}$-stable $A_{\mathrm{c}\ell}$ by enforcing $Z \succ 0$, see *e.g.* Theorem 7.2.5 in [35]. 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 (18) is indeed semi-algebraic for $\gamma > 0$, thus a set of polynomial inequalities in the elements of $\Delta_{A_{\mathrm{c}\ell}}$ of the form $\mathcal{S}$:

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

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 [36].

Proof: [Proof of Lemma III.5] Consider the problem

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

If γ satisfies $h(\delta) = \gamma$ then from the Lemma VI.4 the solution to $\mathcal{P}_a(\gamma)$ can be directly retrieved from the (negated) problem

$$\mathcal{P}_b(\delta): \begin{cases} \underset{\Delta_{A_{\mathrm{c}\ell}} \in \mathbb{R}^{d\times n}}{\mathrm{argmin}} & \underset{x_0, v}{\mathbb{E}} \left[\displaystyle\sum_{k=0}^{\infty} \alpha^k \left(\delta^{-1} w_k^\top w_k - x_k^\top Q_{\mathrm{c}\ell} x_k\right)\right] \\ \text{subject to} & x_{k+1} = \widehat{A}_{\mathrm{c}\ell} x_k + D w_k + v_k, \\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0, \Sigma_v),\ x_0 \sim \mathcal{P}(0, \Sigma_0), \\ & w_k = \Delta_{A_{\mathrm{c}\ell}} x_k. \end{cases} \tag{19}$$

Under the conditions from Proposition III.5 the program $\mathcal{P}_b(\delta)$ can be solved using Dynamic Programming, *e.g.* see chapter 3 from [37], 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 (19), 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 (19)

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

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

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

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.6: Now this apparent link between the solution to a robust LQR problem and a dynamic game is formalized. This is not new, see for example [16], [25], where in the latter[10], the pair (γ, δ) is interpreted via multiplier theory (*cf.* [38], [39]) 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 δ, given γ (see Lemma VI.3).

Recall Definition III.1 and the RLQR problem (3). Let a solution to (3) be denoted by the pair $\big(K^\star(\gamma), \Delta_A^\star(\gamma)\big)$ whereas a solution to (23), 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 (23) to the solution of the robust LQ regulator (3). This proof of Theorem III.6 is split up into a few parts.

Proof: [Proof of Theorem III.6 part (i), (ii) and monotonicity of (iii)] Regarding the monotonicity in (iii), first consider the game (23). 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

$$
f(K,L) = \mathop{\mathbb{E}}_{x_0, v}\left[\sum_{k=0}^{\infty} \alpha^k x_k^\top \left(Q + K^\top R K\right) x_k\right], \tag{20}
$$

$$
g(K,L) = \mathop{\mathbb{E}}_{x_0, v}\left[\sum_{k=0}^{\infty} \alpha^k w_k^\top w_k\right] = \left\langle L^\top L, \Sigma_x \right\rangle, \tag{21}
$$

with $\Sigma_x = \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty} \alpha^k x_k x_k^\top\right]$[11]. Then $\sup_L\{f(K', L) - \delta^{-1} g(K', L)\}$ corresponds to program $\mathcal{P}_2$ from Lemma VI.4 with the map h from (7) 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 (23) are of the form

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

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

respectively, for $f(K,L)$ and $g(K,L)$ defined by (20) and (21).

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.4 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.4 we have $L_1^\star(\gamma, K) = L_2^\star\big(h^{-1}(\gamma), K\big)$. 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.4 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 (23) 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 (23), *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.6, which is again very useful with numerical algorithms in mind. This is done in the spirit of the work by Polderman [41], [42].

Proof: [Proof of Theorem III.6 (iii) cont.] We will first show that $\overline{P}^+(\delta)$[12] is analytic over $[0, \overline{\delta})$, 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

$$
\begin{aligned}
\ell(\delta, P) =& P - Q - \alpha A^\top P \cdots \\
&\cdots \left(I_n + \alpha\left(B R^{-1} B^\top - \delta D D^\top\right) P\right)^{-1} A.
\end{aligned} \tag{22}
$$

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

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

[11] This step relies on the Bounded Convergence Theorem (*cf.* p.57 [40]) 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.,* $\lim_{n\to\infty} \int_{\mathcal{X}} f_n + g_n \mathrm{d}\mu = \int_{\mathcal{X}} \lim_{n\to\infty} f_n \mathrm{d}\mu + \int_{\mathcal{X}} \lim_{n\to\infty} g_n \mathrm{d}\mu$.

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

Δ_P:

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

These steps hinge on geometric series for matrices, and a few linear algebraic identities[13]. Now since we know that $\widetilde{\Lambda}^{-1}A$ is $\sqrt{\alpha}$-stable when $\widetilde{P}$ is $\overline{P}^+(\widetilde{\delta})$, the map Γ must be non-singular (see Lemma 2.3 [41]) for such a point $\big(\widetilde{\delta},\overline{P}^+(\widetilde{\delta})\big)$. Therefore, we can apply the Implicit Function Theorem (*cf.* [43]), which asserts (locally) the existence of an unique C^ω map $P(\delta)$ such that $\ell(\delta,P(\delta))=0$ for all $\delta\in U_{\widetilde{\delta}}\subset\mathbb{R}_{\geq 0}$ plus $P(\widetilde{\delta})=\widetilde{P}$. Since the pair $(\widetilde{\delta},\widetilde{P})$ was arbitrary, up to being a minimal solution, this holds for any pair $(\delta,\overline{P}^+(\delta))$, making $\overline{P}^+(\delta)\in C^\omega\big((0,\overline{\delta})\big)$ since $P(\delta)|_{\delta\in U_{\widetilde{\delta}}}$ are unique (see [44]) and stabilizing by continuity. This implies that $L^\star(\delta)$ is C^ω in δ and by Theorem E.1.4[14]. from [46], so is Σ_x, such that indeed the map $h(\delta)$ is analytic over some bounded interval. Finally, to extend $(0,\overline{\delta})$ to $[0,\overline{\delta})$ observe that $\lim_{\delta\downarrow 0}h(\delta)=0$, which concludes the proof.

Proof: [Proof of Lemma III.8] This follows directly from Theorem III.6 whereas the decomposition follows from any standard proof of Lemma VI.2, *e.g.,* solving the first step in the corresponding Bellman-Isaacs equation (*cf.* [26]).

To prove Lemma III.9 we need one useful property of $\Lambda(\delta)$:

Lemma VI.1 ($\Lambda(\delta)$ is an orientation preserving map)**:** The matrix $\Lambda(\delta)$ has positive eigenvalues and thus $\det\big(\Lambda^{-1}(\delta)\big)>0$.

Proof: The map $\Lambda(\delta)=\big(I_n+\alpha(BR^{-1}B^\top-\delta DD^\top)P(\delta)\big)$ has positive eigenvalues for $\delta\downarrow 0$ since $\lim_{\delta\downarrow 0}\Lambda(\delta)=(I_n+\alpha BR^{-1}B^\top P)$ and any product of (symmetric) positive semi-definite matrices has again positive eigenvalues (although it might fail to remain positive semi-definite). Then recall the fact that $\mathsf{GL}(n,\mathbb{R})$ has two connected components denoted $\mathsf{GL}^+(n,\mathbb{R})$ and $\mathsf{GL}^-(n,\mathbb{R})$ for the orientation *preserving* and *-reversing* maps, respectively. Then the result follows from $\lim_{\delta\downarrow 0}\Lambda(\delta)\in\mathsf{GL}^+(n,\mathbb{R})$ and continuity in δ, *i.e.,* the matrix $\Lambda(\delta)$ cannot leave the set of orientation-preserving non-singular matrices for $\delta\in[0,\overline{\delta})$.

Proof: [Proof of Lemma III.9] We do the proof per item:

(i) The fact that the worst-case closed-loop system can be written as $\big(\Lambda^\star(\delta)\big)^{-1}\widehat{A}$ follows from Lemma VI.2 and Lemma VI.1 or just by direct computation. This also holds for $\gamma\to 0$ since it also holds for the standard LQR closed-loop system[31]. The last part follows from (12), $K^\star(\gamma)$ is always of the form $X\widehat{A}$ for some matrix X. Of course, the intuition is that if your goal is regulation, then once $x_k\in\mathrm{Ker}(A)$ it makes no sense to further inject energy in the system. Therefore, any additive perturbation Δ_A to $\widehat{A}$ must obey $\mathrm{Ker}(\widehat{A})\subseteq W^+\big(\sqrt{\alpha}(\widehat{A}+\Delta_A)\big)$ when $\Sigma_0\succ 0$.

(ii) Lemma VI.1 has several implications. For example, it is known that the worst-case closed-loop system is given by $\Lambda^{-1}(\delta)\widehat{A}$, which has thus the same orientation as $\widehat{A}$. Moreover, it is known that the worst-case drift term is given by $A^\star(\gamma)=\widehat{A}+D\Delta_A^\star(\delta)=(I+\delta\alpha DD^\top P\Lambda^{-1})\widehat{A}$. Also, it follows from equation $(3.4a'')$ in [26] that $P\Lambda^{-1}\succeq 0$, so indeed, now we do have symmetry. So when for example $D=I_n$, we have that the nominal- and worst-case drift have the same orientation. To intuitively see why we speak of orientation-preserving, take the SVD of any $T\in\mathsf{GL}^+(n,\mathbb{R})$ which is $T=U\Sigma V^\top$, where both U and V are rotation matrices, while Σ is a positive scaling matrix. Then $T\widehat{A}$ will be a rotated and scaled version of $\widehat{A}$, no other operations, like mirroring, occur. Note that actually, the scaling matrix T is an element of $\mathcal{S}^n_{++}$. When $\widehat{A}$ is not full-rank, we can without loss of generality take just the automorphic part.

(iii) We know that $A^\star(\gamma)$ is of the form $(I_n+\alpha\delta P\Lambda^{-1})\widehat{A}=T\widehat{A}$, $T\in\mathcal{S}^n_{++}$. This means that $\lambda_{\min}(T)\geq 1$ or $\lambda_{\min}(T)>1$ *a.s.* when $P\succ 0$. Now embed $\widehat{A}$ into n^2 and such that $\mathrm{vec}\big(A^\star(\gamma)\big)=(I_n\otimes T)\mathrm{vec}(\widehat{A})$. The spectrum and symmetry of T are preserved in $(I_n\otimes T)$ such that we can appeal to inequalities of the form $\lambda_{\min}(Y)\|x\|_2\leq\|Yx\|_2\leq\lambda_{\max}(Y)\|x\|_2$, $Y\in\mathcal{S}^{n^2}_{++}$. Hence, the transformation will make any vector grow in 2-norm. The results follows from the element-wise interpretation of the Frobenius-norm. Regarding the decomposition (15), using the identity $\big(I+(I-P)^{-1}P\big)=(I-P)^{-1}$ we can write $B^\star(\gamma)$ as $\big(I_n-\alpha\delta P(\delta)\big)^{-1}\widehat{B}$. Then the result follows from $\big(\delta^{-1}I_n-\alpha P(\delta)\big)\succ 0$, symmetry of P and a similar line of arguments as above.

B. Supporting Material

1) Dynamic Game Theory: It should be highlighted that the link between dynamic game theory and robust control is well studied, see [16], [26] for an accessible and illuminating introduction.

We first introduce the concept of a dynamic game (*cf.* [26], [47]). Define a real-valued map g by $g(x,u,w)=x^\top Qx+u^\top Ru-\delta^{-1}w^\top w$ and consider the stochastic (discounted)

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

[14]Effectively, by the results from Polderman [45]

two-player zero-sum *dynamic game*

$$\begin{aligned}\inf_{\{\mu_k\}_{k\in\mathbb{N}}} \sup_{\{\nu_k\}_{k\in\mathbb{N}}} \quad & \mathbb{E}_{x_0,v}\left[\sum_{k=0}^{\infty}\alpha^k g(x_k,u_k,w_k)\right], \\ \text{s.t.} \quad & x_{k+1} = Ax_k + Bu_k + Dw_k + v_k, \\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0,\Sigma_v), \quad x_0 \sim \mathcal{P}(0,\Sigma_0), \\ & u_k = \mu_k(x_k), \quad w_k = \nu_k(x_k).\end{aligned} \tag{23}$$

Here, the parameter $\delta \in \mathbb{R}_{\geq 0}$ penalizes the input of the μ-player, which reduces its ability to destabilize the system, and $D \in \mathbb{R}^{n\times d}$ determines how the state dynamics are affected by the input of this ν-player. Note that this game is "*diagonal*"[15] in the sense that there are no cross-terms in the cost, thus the program largely relies on the single parameter δ.

This next Lemma summarizes the key results we need regarding the dynamic game

Lemma VI.2 (*cf.* chapter 3 from [26] for the undiscounted deterministic case)**:** Given a game (23) 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 $\delta \in \mathbb{R}_{\geq 0}$ satisfies $(\delta^{-1}I_d - \alpha D^\top P D) \succ 0$[16], where P is the minimal[17] positive semi-definite solution to the Generalized Algebraic Riccati Equation (GARE):

$$\begin{aligned} P =& Q + \alpha A^\top P \Lambda^{-1} A, \\ \Lambda =& \left(I_n + \alpha\left(BR^{-1}B^\top - \delta DD^\top\right)P\right), \end{aligned} \tag{24}$$

then the optimal[18] 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

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

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

2) Computational Remarks, Given γ, Find δ: The main computational question is twofold, given a $\gamma \in \mathbb{R}_{\geq 0}$, *(i)* does there exist a $\delta \in \mathbb{R}_{\geq 0} : h(\delta) = \gamma$ and *(ii)*, if so, how to find it? Regarding question *(i)*, by monotonicity it suffices to find a upper bounding $\overline{\gamma}$ and show that $\gamma \leq \overline{\gamma}$ (see [27, sec. 3-4-2] for limiting behaviour of the map h, which can be finite).

This can be done by finding an upper bound to $\overline{\delta}$. The idea is that since $P(\delta) \succeq P(0)$ the solution to

$$\begin{aligned} \sup_{\delta\in\mathbb{R}_{\geq 0}} \quad & \delta \\ \text{subject to} \quad & \delta^{-1}I_d - \alpha D^\top P(0) D \succeq 0 \end{aligned} \tag{25}$$

[15] 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 $S \succeq 0$.

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

[17] In the terminology of p.81 ch.3 [26], 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.

[18] Not a general saddle-point (see *e.g.* [48])

upper bounds $\overline{\delta}$. Recall that with $P(0)$ we mean the stabilizing solution to the standard discounted Algebraic Riccati Equation. Since $(\sqrt{\alpha}A, B, C)$ should be a minimal realization (see Lemma VI.2), $P(0)$ exists, such that the solution to (25) is given by $\delta^\star = \|\alpha D^\top P(0) D\|_2^{-1}$. Of course, for a meaningful bound we must assume that $DP(0) \neq 0$.

The crux is, we can come arbitrary close to $\overline{\delta}$ by using bisection, which also yields a bound on γ. Then by applying bisection again, we can solve the problem, or conclude infeasibility. So, to solve any of our robust LQR problems we have a slow, yet tractable, procedure as summarized in Figure 8.

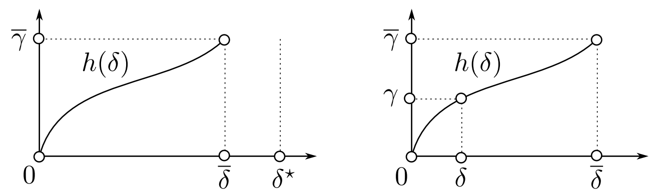


Fig. 8: Let us be given a Robust LQR problem for some γ. First (1), compute $\delta^\star$ from (25) and find $\overline{\delta}$ using any algorithm similar to bisection. As a byproduct, $\overline{\gamma}$ is given such that feasibility of γ can be readily checked. Note that in practice one rather wants to find $\overline{\delta}$ from below, *i.e.* find some δ being ε-close to $\overline{\delta}$ (denoted $\overline{\delta}_\varepsilon$), such that $\overline{\delta} - \overline{\delta}_\varepsilon = \varepsilon > 0$, guaranteeing feasibility of the corresponding dynamic game related to $\overline{\delta}_\varepsilon$, leading to the bound $\overline{\gamma}_\varepsilon \leq \overline{\gamma}$. Then secondly (2), using $\{0, \overline{\delta}_\varepsilon\}$ as starting pair, one can apply any algorithm similar to bisection to find $\delta : h(\delta) = \gamma$ for $\gamma \leq \overline{\gamma}_\varepsilon$.

Regarding question *(ii)*, as already mentioned, the properties of the map h allow for bisection algorithms indeed, but when one has more insights in the shape of h its image, convergence *can* be much faster.

Lemma VI.3 (Finding δ)**:** Given a desired γ and assume it is feasible in the sense of Theorem III.6. Let the (local) Lipschitz constant of the map h be $L > 0$ on $[0, \delta)$ and select $\beta \leq L^{-1}$. Then, the algorithm

$$\delta_{k+1} = \delta_k + \beta\big(\gamma - h(\delta_k)\big), \quad \delta_0 = 0, \tag{26}$$

converges to $\delta : h(\delta) = \gamma$ at a linear rate proportional to the estimation error of L.

Proof: [Proof of Lemma VI.3] Consider the algorithm

$$\delta_{k+1} = \delta_k + \beta_k\big(\gamma - h(\delta_k)\big), \quad \delta_0 = 0. \tag{27}$$

To find a suitable sequence of stepsizes $\{\beta_k\}_{k\in\mathbb{N}}$ define the error $e_k := \delta - \delta_k$ and consider the Lyapunov candidate $V_k = e_k^2$. Then we need to find β_k such that $V_k - V_{k+1} > 0$ for non-zero errors. It can be easily seen that a satisfactory constraint on β_k is

$$\beta_k < \frac{2(\delta - \delta_k)}{\gamma - h(\delta_k)}.$$

Since the map h is (locally) smooth, it is definitely locally Lipschitz, *i.e.,* we have for some constant $L > 0$

$$|h(\delta_2) - h(\delta_1)| \leq L|\delta_2 - \delta_1|, \quad \delta_1, \delta_2 \in [0, \overline{\delta}). \tag{28}$$

Therefore, by (28) and monotonicity of h, the constraint on β_k can be simplified to $\beta_k < 2/L \ \forall k$. Therefore, simply setting $\beta_k = L^{-1}$ works. Note that we have not yet provided a method to compute L, thus the constant must estimated, denote this by $\widehat{L}$ for which $\widehat{L} \geq L$ must hold. The error dynamics are given by $e_{k+1} = e_k - \widehat{L}^{-1}\big(\gamma - h(\delta_k)\big) = (1-\varepsilon)e_k$, for some $\varepsilon \in (0,1]$ such that, the cruder $\widehat{L}$ is, the smaller ε and thus the slower $e_{k+1} \to 0$.

We do emphasize that estimation of the Lipschitz constant is critical the closer $\delta : h(\delta) = \gamma$ is to $\overline{\delta}$.

3) Auxiliary Tools: 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.4 (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).[19]

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

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

Adding the two sides of the above inequalities yields

$$\begin{aligned} (\gamma_2^{-1} - \gamma_1^{-1}) g\big(x_2^\star(\gamma_2)\big) \leq (\gamma_2^{-1} - \gamma_1^{-1}) g\big(x_2^\star(\gamma_1)\big) \iff \\ g\big(x_2^\star(\gamma_2)\big) \leq g\big(x_2^\star(\gamma_1)\big) \end{aligned}$$

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 $\gamma \geq 0$. 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 γ 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 $g(\widetilde{x}_1) \leq g\big(x_2^\star(\gamma)\big)$. 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).

Lemma VI.5 (Monotonic Factor)**:** Given $\delta_1 \geq \delta_2$, both in $(0, \overline{\delta})$ and corresponding to a feasible game. Then

$$\delta_1^{-1}(\delta_1^{-1} I_d - \alpha D^\top P(\delta_1) D)^{-1} \succeq \delta_2^{-1}(\delta_2^{-1} I_d - \alpha D^\top P(\delta_2) D)^{-1}.$$

Proof: [Proof of Lemma VI.5] Since $\delta_1 \geq \delta_2 > 0$ we have $P(\delta_1) \succeq P(\delta_2)$ such that

$$\begin{aligned} (I_d - \delta_2 \alpha D^\top P(\delta_2) D) \succeq (I_d - \delta_1 \alpha D^\top P(\delta_1) D) \iff \\ \delta_1^{-1}(\delta_1^{-1} I_d - \alpha D^\top P(\delta_1) D)^{-1} \succeq \delta_2^{-1}(\delta_2^{-1} I_d - \alpha D^\top P(\delta_2) D)^{-1}. \end{aligned}$$

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 μ 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^∞ 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^∞ Methods*. Springer, 2000.

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

[18] Y. Abbasi-Yadkori and C. Szepesvári, “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] W. Jongeneel, T. Summers, and P. Mohajerin Esfahani, “Robust linear quadratic regulator: Exact tractable reformulation,” in *2019 IEEE Conference on Decision and Control (CDC)*, 2019.

[24] R Abraham, J. Marsden, and T Ratiu, *Manifolds, Tensor analysis, and Applications, Second Edition*, ser. Applied Mathematical Sciences. Springer, 1988.

[25] 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.

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

[27] W. Jongeneel, “Controlling the unknown: A game theoretic perspective,” Master’s thesis, Systems & Control, Delft University of Technology, 2019. [Online]. Available: {wjongeneel.nl/pub/mscThesis_wjongeneel.pdf}.

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

[29] M. Yoon, V. A. Ugrinovskii, and I. R. Petersen, “On the worst-case disturbance of minimax optimal control,” in *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, vol. 1, 2002, 604–609 vol.1.

[30] N. Kuiper and J. Robbin, “Topological classification of linear endomorphisms,” *Inventiones math. Springer Verlag*, vol. 19, pp. 83–106, 1973.

[31] J. Polderman, “On the necessity of identifying the true parameter in adaptive lq control,” *Systems & Control Letters.*, no. 8, pp. 87–91, 1986.

[32] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, “Regret bounds for robust adaptive control of the linear quadratic regulator,” in *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., Curran Associates, Inc., 2018, pp. 4188–4197.

[33] H. Mania, S. Tu, and B. Recht, *Certainty equivalence is efficient for linear quadratic control*, 2019. arXiv: 1902.07826.

[34] I. Houtzager, B. Kulcsar, J.-W. van Wingerden, and M. Verhaegen, “System identification for the control of wind turbine systems,” in *The Science of Making Torque from Wind*, ser. TORQUE 2010, Crete, Greece, 2010, pp. 675–680.

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

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

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

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

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

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

[41] 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.

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

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

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

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

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

[47] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. SIAM, 1999.

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